

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 17, 2002, 08:24:35 ; Search time 1208.78 Seconds
(without alignments)
256.813 Million cell updates/sec

Title: US-09-780-113D-15
Perfect score: 23
Sequence: 1 accacgaactgaacgaacactt 23

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estba:*
2: em_estbu:*
3: em_estlu:*
4: em_estlu:*
5: em_estlv:*
6: em_estlp:*
7: em_estro:*
8: em_hlc:*
9: qb_estli:*
10: qb_estli2:*
11: qb_hlc:*
12: qb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18.2	79.1	530 12	AA057613 HS-5097-B
2	18.2	79.1	914 10	BC165582 602343735
3	17.8	77.4	375 9	AW478086 18502 MAR
4	17.8	77.4	697 10	B1875875 116407.Y
5	17.4	75.7	511 10	BM354761 rr16e10.Y
6	17.4	75.7	401 10	BM343863 rr43h05.Y
7	17.2	74.8	246 10	W06581 72484 MVAR4
8	17.2	74.8	455 9	AI621053 ts76d07.X
9	17.2	74.8	475 10	B1345371 374175 MA
10	17.2	74.8	574 12	BM622013 1007115B0
11	17.2	74.8	632 10	B1886608 ZF637-1-0
12	17.2	74.8	997 12	CNS05NKN
13	17.2	74.8	1302 10	BE9652881
14	17.2	73.9	476 10	BR652364
15	17.2	73.9	520 10	BR652626
16	16.8	73.0	325 9	BE109045
17	16.8	73.0	346 9	AU052055

C 18	16.8	73.0	386 10	BF332831	BF332831 CMT_R1076
C 19	16.8	73.0	594 9	A1940637	A1940637 IL4-S1002
C 20	16.8	73.0	601 9	A1940638	A1940638 IL4-S1002
C 21	16.8	73.0	654 9	AM057195	AM057195 CA02408.Y
C 22	16.8	73.0	896 10	BI086804	BI086804 G0285041.3
C 23	16.8	73.0	947 10	BF204564	BF204564 60188174
C 24	16.8	73.0	1409 10	BF237879	BF237879 60188174
C 25	16.6	72.2	297 9	BB369390	BB369390 BS303930
C 26	16.6	72.2	309 9	BB308471	BB308471 BS308471
C 27	16.6	72.2	327 9	BB376297	BB376297 BS376297
C 28	16.6	72.2	348 10	BF765798	BF765798 IL2-OS004
C 29	16.6	72.2	435 12	A2058336	A2058336 RPT-2-4
C 30	16.6	72.2	460 10	BF036844	BF036844 601459173
C 31	16.6	72.2	508 9	AA922286	AA922286 om4h008.S
C 32	16.6	72.2	522 12	AO550432	AO550432 RPT-11.4
C 33	16.6	72.2	532 10	BF768721	BF768721 QV4-11000
C 34	16.6	72.2	545 12	TA391F030	TA391F030 IL4-98462.1
C 35	16.6	72.2	551 10	BE450880	BE450880 EST-101767
C 36	16.6	72.2	555 12	BM403758	BM403758 AG-NO-151
C 37	16.6	72.2	567 12	B68429	B68429 C11-HSV-202
C 38	16.6	72.2	574 9	A0145954	A0145954 A0145954
C 39	16.6	72.2	594 12	A2093474	A2093474 RPT-2-4
C 40	16.6	72.2	595 9	AM934499	AM934499 EST1354105
C 41	16.6	72.2	622 12	AO984457	AO984457 RPT-2-3
C 42	16.6	72.2	632 10	BM426935	BM426935 PM120-PK0
C 43	16.6	72.2	637 12	A2115664	A2115664 RPT-2-1
C 44	16.6	72.2	648 12	A2574567	A2574567 340P610
C 45	16.6	72.2	665 12	A2222176	A2222176 RPT-2-3.6

ALIGNMENTS

RESULT 1
AA057613/c
LOCUS
DEFINITION
AA057613
ACCSSION
AA057613.1 GI:4636253
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Fukuyama; Melacoe; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 530)
Keller, A., Shaker, R., Fulong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.

TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPT-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.hsc.washington.edu
plate: 673 row: P column: 13
Seq primer: 77
Class: BAC ends
High quality sequence stop: 530.
Location/Qualifiers
1..530
/organism="Homo sapiens"

FEATURES
source

AUTHORS
 Sugano, S., Kawakami, K., Johnson, S., Li, F., Marra, M., Eddy, S.,
 Miller, L., Clifton, S., Allen, M., Gibbons, M., Jost, S., Kucaba, T.,
 Martin, J., Pape, D., Stegmeier, M., Underwood, K., Theising, B., Rittler,
 E., Bowers, Y., Wylie, T., Waterston, R. and Wilson, R.
JOURNAL
 Unpublished (1999)
COMMENT
 Contact: S.L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 Library constructed by Dr. Sumio Sugano and Dr. Kohji Kawakami DNA
 sequencing by: Washington University Genome Sequencing Center (clone
 distribution information can be found through the J.M.A.G.E.
 Consortium/JMNC, send email to: infoimage.jml.gov
 High quality sequence stop: 520.

FEATURES
Source
 1..697
 /organism="Danio rerio"
 /strain="AB"
 /db_xref="taxon:7955"
 /clone="3818460"
 /clone_lib="Sugano Kawakami zebrafish DNA"
 /sex="mixed (one male and one female, including
 unfertilized eggs)"
 /dev_stage="adult"
 /lab_host="DH10B (phage resistant)"
 /note="Vector: pME18S-FL3; site_1: DraIII (CAGCGTGTG);
 site_2: DraIII (CAGCATGTG); 1st strand cDNA was primed
 with an oligo(dT) primer (ATGTGCGCTTTTCTTTTCTTTT);
 double-stranded cDNA was ligated to a DraIII adaptor
 (TTGTGCGTGTGTG), digested and cloned into distinct DraIII
 sites of the pME18S-FL3 vector (5' site CAGCGTGTG, 3' site
 CAGCATGTG). XhoI should be used to isolate the cDNA
 insert. Size selection was performed to exclude fragments
 <1.5kb. Library constructed by Dr. Sumio Sugano
 (University of Tokyo Institute of Medical Science) and
 kindly donated by Dr. Kohji Kawakami. Custom primers for
 sequencing: 5' end primer CTCTGCTCTTAACACCTCCG and 3' end
 primer CGACCTGCACCTGCACCA."

BASE COUNT
 213 a 155 c 183 g 146 t

ORIGIN

Query Match
 Best local similarity 77.4%; Score 17.8; DB 10; Length 697;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

LOCUS
 BM354761 401 bp mRNA linear EST 23-JAN-2002

DEFINITION
 rib6l0.y1 Globodera rostochiensis J2 pCDNA11 Smart v1 Globodera
 rostochiensis cDNA 5' similar to TR:Q18231 Q18231 SIMILAR TO 40S
 RIBOSOMAL PROTEIN S30/UBIQUITIN-LIKE PROTEIN. [1] ; mRNA sequence.

ACCESSION
 BM354761

VERSION
 BM354761.1 GI:18089392

KEYWORDS
 EST.

SOURCE
 Globodera rostochiensis.

ORGANISM
 Globodera rostochiensis
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heteroderae; Heteroderidae; Globodera.

REFERENCE
 1 (bases 1 to 401)
 McCarter, J., Clifton, S., Chiappelli, B., Pape, D., Martin, J., Wylie, T.,
 Dante, M., Marra, M., Miller, L., Kucaba, T., Theising, B., Bowers, Y.,
 Gibbons, M., Rittler, E., Bennett, J., Franklin, C., Tsagarisvilli, R.,
 Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Stegmeier,
 M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,
 Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and

AUTHORS
 Wilson, R.
JOURNAL
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
COMMENT
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 The library was contributed by Dr. Geert Smant of the Laboratory of
 Nematology at Wageningen University, Wageningen,
 Netherlands (geert.smant@wur.nl). DNA sequencing by:
 Washington University Genome Sequencing Center
 -- q primer: -40RP from Gibco
 High quality sequence stop: 393.

FEATURES
Source
 1..401
 /organism="Globodera rostochiensis"
 /db_xref="taxon:31243"
 /clone_lib="Globodera rostochiensis J2 pCDNA11 Smart v1"
 /dev_stage="J2"
 /lab_host="DH10B"
 /note="Vector: pCDNA11 (Invitrogen); site_1: bsrXI;
 site_2: EcoRI; The library was donated for sequencing by
 Geert Smant from Wageningen University, Laboratory of
 Nematology, The Netherlands."

BASE COUNT
 200 a 58 c 80 g 62 t 1 others

ORIGIN

Query Match
 Best local similarity 75.7%; Score 17.4; DB 10; Length 401;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

LOCUS
 BM433863 511 bp mRNA linear EST 23-JAN-2002

DEFINITION
 r743h05.y1 Globodera rostochiensis J2 pCDNA11 Smart v1 Globodera
 rostochiensis cDNA 5' similar to TR:Q9XX24 Q9XX24 HB10111N/5'AA
 FUSION PROTEIN. ; mRNA sequence.

ACCESSION
 BM433863

VERSION
 BM433863.1 GI:18080778

KEYWORDS
 EST.

SOURCE
 Globodera rostochiensis.

ORGANISM
 Globodera rostochiensis
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heteroderae; Heteroderidae; Globodera.

REFERENCE
 1 (bases 1 to 511)
 McCarter, J., Clifton, S., Chiappelli, B., Pape, D., Martin, J., Wylie, T.,
 Dante, M., Marra, M., Miller, L., Kucaba, T., Theising, B., Bowers, Y.,
 Gibbons, M., Rittler, E., Bennett, J., Franklin, C., Tsagarisvilli, R.,
 Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Stegmeier,
 M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S.,
 Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and
 Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 The library was contributed by Dr. Geert Smant of the Laboratory of
 Nematology at Wageningen University, Wageningen,
 Netherlands (geert.smant@wur.nl). DNA sequencing by:
 Washington University Genome Sequencing Center

Seq primer: -40bp from 5' end
High quality sequence slope: 479.

FEATURES

Location/Qualifiers

1..511

/organism="Glossosoma rostrichionensis"

/db_xref="taxon:31243"

/clone_lib="Glossosoma rostrichionensis J2 pGEM11 Smart v1"

/vec_state="32"

/lab_host="DH10B"

/note="Vector: pGEM11 (Invitrogen); Site_1: 881X1; Site_2: EcoRI; The library was donated for sequencing by Dr. Stuart from Wageningen University, Laboratory of Nematology, The Netherlands."

BASE COUNT

143 a 106 c 144 g 118 t

QUERY MATCH

Best local similarity 75.7%; Score 17.4; DB 10; Length 511;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGACGACGACGACGACCTT 23

DB 401 GGTCTGACGACGACCTT 283

RESULT 7

W06581 246 bp mRNA linear EST 12-AUG-1996

LOCUS

T2484 MYA14 bloodstream form of seropome WMA11.1 Trypanosoma

brucei Rhodesiense cDNA 5', mRNA sequence.

ACCESSION

W06581

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

COMMENT

CONTACT

TELEPHONE

TELEFAX

EMAIL

Seq primer: 13 primer.

Location/Qualifiers

1..246

/organism="Trypanosoma brucei Rhodesiense"

/db_xref="taxon:31286"

/clone_lib="MYA14 bloodstream form of seropome WMA11.1"

/note="Vector: Lambda ZAP 11 (Stratagene); Site_1: EcoRI; Site_2: XhoI; The mRNA was purified from a cloned population of bloodstream trypanosomes repressing the MYA14 metacyclic variant surface glycoprotein (VSG). A unidirectional oligo dT-primed EcoRI/XhoI cDNA library was constructed in Lambda ZAP 11 (Stratagene)."

BASE COUNT

78 a 26 c 83 g 59 t

ORIGIN

Query Match

Best local similarity 74.8%; Score 17.2; DB 10; Length 246;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACGACGACGACGACCTT 23

DB 211 CCTTACTGACGACGACCTT 190

RESULT 8

A1621053/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

COMMENT

JOURNAL

CONTACT

EMAIL

Tissue procurement: Christopher A. Muskulok, M.D., Ph.D., Michigan

R. Emmert-Buck, M.D., Ph.D.

cDNA Library preparation: M. Kenton Soares, Ph.D., M. Fatima

Bernaldo, Ph.D.

cDNA Library Arrayed by: Greg Leaman, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be

found through the J.M.A.G.E. Consortium/J.M.R. at:

www.bio.1ind.gov/brp/image/image.html

Insert length: 612 Std Error: 0.00

Seq primer: -40bp from 5' end

High quality sequence slope: 449

POLYA-No.

Location/Qualifiers

1..455

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:2237197"

/clone_lib="NCI_CGAP_G66"

/tissue_type="Pooled germ cell tumors"

/note="Vector: p7730-Pac (pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA

from the normalized library NCI-CGAP G64 was prepared, and

as circles were made in vitro. Following BAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was FPK-amplified cDNAs from a pool

of 5,000 clones made from the same library (clones

1257095-1258631, 1469064-1470983, and 1475922-1476743).

Subtraction by Bento Soares and M. Fatima Bernaldo.

BASE COUNT

76 a 143 c 148 g 88 t

ORIGIN

Query Match

Best local similarity 74.8%; Score 17.2; DB 9; Length 455;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACGACGACGACGACCTT 23

DB 239 CCACGACGACGACGACCTT 218

RESULT 9

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 475)

161445471 475 bp mRNA linear EST 30-JUL-2001

B1345371

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 475)


```

BASE COUNT      172 a   144 c   154 g   162 t
ORIGIN
Query Match
Best Local Similarity 74.8%; Score 17.2; DB 10; Length 612;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY      2 CCACGACTGACGACGACCTTT 23
      11111111111111111111
      505 CCACGACGACGACGACCTCTT 484

RESULT 12
LOCUS      CNS05NMK
DEFINITION Tetradodon nigroviridis genome survey sequence T7 end of clone
           019601 of library B from Tetradodon nigroviridis, genomic survey
           sequence.
ACCESSION  AF345488.1 GI:8239258
VERSION     AF345488
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetradodon nigroviridis.
ORGANISM    Tetradodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
            Tetraodontidae; Tetradodon.
REFERENCE   1 (bases 1 to 997)
AUTHORS     Koest-Crolius,H., Jallion,O., Dasilva,C., Filames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernol,A. and
            Weissenbach,J.
TITLE        Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetradodon nigroviridis
            Unpublished
            2 (bases 1 to 997)
            Koest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernol,A., Filames,C., Wincker,P., Brothier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
            Human gene number estimate provided by genome wide analysis using
            Tetradodon nigroviridis DNA sequence
            Unpublished
            3 (bases 1 to 997)
            Genoscope.
            Direct Submission
            Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
            This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetradodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetradodon.
FEATURES
    source
        1..997
        Location/Qualifiers
            /organism="Tetradodon nigroviridis"
            /db_xref="taxon:99883"
            /clone_lib="B"
            /note="Genoscope sequence ID : COAR019AD01C1-end : T7"

BASE COUNT      265 a   168 c   240 g   303 t
ORIGIN
Query Match
Best Local Similarity 74.8%; Score 17.2; DB 12; Length 997;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY      2 CCACGACTGACGACGACCTTT 23
      11111111111111111111
      781 CCACGACGACGACGACCTGT 760

RESULT 13
LOCUS      BF652881
DEFINITION BF652881 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3856229 3',
            mRNA sequence.
            1302 bp mRNA linear EST 14-DEC-2000
            BF652881
            mRNA sequence.

ACCESSION  BF652881.2 GI:11766244
VERSION     BF652881.2
KEYWORDS    EST.
SOURCE      human.
            Homo sapiens
            Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 1302)
AUTHORS     NIH-MGC http://mhc.nhl.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            On Oct 3, 2000 this sequence version replaced g1:10575685.
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cygnus-remail.nih.gov
            Tissue Procurement: DCTB/DBP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The L.M.A.G.E. Consortium (LML)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the L.M.A.G.E. Consortium/LML at:
            http://image.llnl.gov
            plate: L1CM571 row: 3 column: 06
            High quality sequence stop: 177.
FEATURES
    source
        1..1302
        Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="NIH_MGC_66"
            /clone="IMAGE:3856229"
            /tissue_type="adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: ovary; Vector: pCMV-Spork6; Site 1: M11;
            Site 2: Salt; Cloned unidirectionally. Primer: 5' to 3'.
            Average insert size 1.8 kb. Library constructed by Life
            Technologies."
            Technology:
            273 a   345 c   317 g   361 t   6 others

BASE COUNT      273 a   345 c   317 g   361 t   6 others
ORIGIN
Query Match
Best Local Similarity 74.8%; Score 17.2; DB 10; Length 1402;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY      2 CCACGACGACGACGACCTTT 23
      11111111111111111111
      172 CCACGACGACGACGACCTTT 151

RESULT 14
LOCUS      BF652464
DEFINITION BF652464 MARC 380V Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BF652464
VERSION     BF652464.1 GI:11917494
KEYWORDS    EST.
SOURCE      cow.
            Bos taurus
            Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
            1 (bases 1 to 476)
            Smith,T.P., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,K.L.,
            Casas,E., Gray,J.E., White,J., Chou,J., Fahrenkrug,S.C., Bennett,
            G.L., Heaton,M.P., Jaegerfeld,W.W., Rohrer,G.A., Chitko-McKorn,C.G.,
            Porter,G., Holt,L., Karayancheva,S., Liang,F., Quackenbush,J. and
            Keefe,J.W.
            Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
            Genome Res. 11 (4), 626-630 (2001)
            Contact: Smith TPL
            USDA, ARS US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366

```

Email: smith@mail.marv.usda.gov
 Single pass sequencing. Bases called and all trimmed with phred
 v0.960904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.

PCR PRIMERS

FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCGACATGACGACG
 Plate: 65 row: F column: 12
 Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers

FEATURES

Source

1..476
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="BARC 5ROV"
 /tissue_type="pooled"
 /lab_host="DH10b"
 /note="Vector: pCMV SPORI6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 BASE COUNT 95 a 105 c 131 g 145 t
 ORIGIN

Query Match 73.9%; Score 17; Dh 10; Length 476;
 Best local Similarity 100.0%; Pred. No. 1.2e+04;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CACGACTGACACGACAC 19
 11111111111111111111
 DB 42H CACGACTGACACGACAC 312

RESULT 15
 HG692626/c 520 bp mRNA linear EST 02-MAY-2001
 LOCUS 342393 BARC 5ROV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION HG692626
 ACCESSION HG692626.1 GI:13944446
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Cow.
 Bos taurus
 Eukaryota; Metazoa; Chordata; Gnathia; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 520)
 Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and
 Wells,K.D.
 Mapping of Expressed Sequence Tags from a normalized bovine mammary
 gland cDNA library
 Unpublished (2000)
 Contact: Sonstegard TS
 USDA, ARS, Beltsville Agricultural Research Center
 Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
 Tel: 301 504 8416
 Fax: 301 504 8414
 Email: tads@psi.barc.usda.gov

TITLE
 JOURNAL
 COMMENT
 Single pass sequencing. Bases called and all trimmed with phred
 v0.960904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.

FEATURES

Source

1..520
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="BARC 5ROV"
 /tissue_type="pooled"
 /lab_host="DH10b"
 /note="Vector: pCMV SPORI6; Site_1: XbaI; Site_2: XhoI;"

Library made from pooled mRNA isolated from mammary
 tissues at eight physiological, developmental, and tissue-
 states.

BASE COUNT 116 a 135 c 140 g 129 t
 ORIGIN

Query Match 73.9%; Score 17; Dh 10; Length 520;
 Best local Similarity 100.0%; Pred. No. 1.2e+04;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CACGACTGACACGACAC 19
 11111111111111111111
 DB 293 CACGACTGACACGACAC 277

Search completed: October 17, 2002, 10:04:01
 Job time : 1217.78 secs

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 3; SEQ ID 6416; 2537pp + CD ROM; English.
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo- or primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 574 BP; 118 A; 144 C; 162 G; 143 T; 7 other;
Query Match 72.2%; Score 16.6; DB 22; Length 574;
Best Local Similarity 82.6%; Pred. No. 79;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCGACGACTGAGCAGCAGCCTT 23
||||| | | | | | | | | |
DB 312 ACCACGCCGCGATCAGCAGCCTCT 290
RESULT 2
AAV69441
ID AAV69441 standard; cDNA; 736 BP.
XX
AC AAV69441;
XX
DE 15-MAR-1999 (first entry)
XX
DE Banana fruit ripening-related clone U-13 cDNA.
XX
KM Fruit ripening; banana; modulator; tissue senescence; crop; plant;
KW triploid; plant breeding; ss.
XX
CS Musa acuminata.
XX
EN WO9853085-A1.
XX
PD 26-NOV-1998.
XX
PF 05-MAY-1998; 98WO-GB01297.
PR 20-MAY-1997; 97GB-0010370.
PA (ZENE) ZENECA LTD.
XX
PI Bird CR, Medina-Suarez RDJ, Seymour GB;
XX
DR WPI; 1999-059745/05.
PT New method of modulating fruit ripening or tissue senescence
PT characteristics of Musa plants - by introducing DNA sequences,

PT useful in processes for modifying plant/fruit ripening
PT characteristics
XX
XX Claim 1; Page 19; 78pp; English.
PS
XX
CC AAV69440-V69512 are cDNA sequences which are used in a method to
CC modulate the fruit ripening or tissue senescence characteristics of Musa
CC acuminata (banana) plants. The method provides a recombinant way of
CC modulating ripening/senescence characteristics of bananas, which are a
CC globally important crop. In particular, the method can modulate such
CC characteristics in the dessert banana, which due to its triploid nature
CC is unaffected by conventional plant breeding techniques.
XX
SQ Sequence 736 BP; 186 A; 190 C; 196 G; 146 T; 18 other;
Query Match 72.2%; Score 16.6; DB 20; Length 736;
Best Local Similarity 82.6%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCGACGACTGAGCAGCAGCCTT 23
||||| | | | | | | | | |
DB 281 ACCTCGACTGAGCAGCAGCCTT 303
RESULT 3
AAV69443
ID AAV69443 standard; cDNA; 736 BP.
XX
AC AAV69443;
XX
DE 15-MAR-1999 (first entry)
XX
DE Banana fruit ripening-related clone U-105 cDNA.
XX
KM Fruit ripening; banana; modulator; tissue senescence; crop; plant;
KW triploid; plant breeding; ss.
XX
CS Musa acuminata.
XX
EN WO9853085-A1.
XX
PD 26-NOV-1998.
XX
PF 05-MAY-1998; 98WO-GB01297.
PR 20-MAY-1997; 97GB-0010370.
PA (ZENE) ZENECA LTD.
XX
PI Bird CR, Medina-Suarez RDJ, Seymour GB;
XX
DR WPI; 1999-059745/05.
PT New method of modulating fruit ripening or tissue senescence
PT characteristics of Musa plants - by introducing DNA sequences,
PT useful in processes for modifying plant/fruit ripening
PT characteristics
XX
PS Claim 1; Page 20; 78pp; English.
XX
CC AAV69440-V69512 are cDNA sequences which are used in a method to
CC modulate the fruit ripening or tissue senescence characteristics of Musa
CC acuminata (banana) plants. The method provides a recombinant way of
CC modulating ripening/senescence characteristics of bananas, which are a
CC globally important crop. In particular, the method can modulate such
CC characteristics in the dessert banana, which due to its triploid nature
CC is unaffected by conventional plant breeding techniques.
XX
SQ Sequence 736 BP; 189 A; 193 C; 202 G; 138 T; 14 other;
Query Match 72.2%; Score 16.6; DB 20; Length 736;
Best Local Similarity 82.6%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCAGACTGAGCAGCACCCTTT 23
 ||| ||||| | ||||| |||
 DB 284 ACCTCGACTGGGAAAGCACCCTT 306

RESULT 4

AAV69442
 ID AAV69442 standard; cDNA: 787 BP.

AC AAV69442;

DT 15-MAR-1999 (first entry)

DE Banana fruit ripening-related clone U-23 cDNA.

KW Fruit ripening; banana; modulator; tissue senescence; crop; plant;

KM triploid; plant breeding; ss.

OS Musa acuminata.

PN WO9853085-A1.

PD 26-NOV-1998.

PF 05-MAY-1998; 98WO-GB01297.

PR 20-MAY-1997; 97GB-0010370.

PA (ZENE) ZENECA LTD.

PI Bird CR, Medina-Suarez RDU, Seymour GB;

DR WPI; 1999-059745/05.

PT New method of modulating fruit ripening or tissue senescence

PT characteristics of Musa plants - by introducing DNA sequences,

PT useful in processes for modifying plant/fruit ripening

PT characteristics

PS Claim 1; Page 19-20; 78pp; English.

CC AAV69440-V69512 are cDNA sequences which are used in a method to

CC modulate the fruit ripening or tissue senescence characteristics of Musa

CC acuminata (banana) plants. The method provides a recombinant way of

CC modulating ripening/senescence characteristics of bananas, which are a

CC globally important crop. In particular, the method can modulate such

CC characteristics in the desert banana, which due to its triploid nature

CC is unaffected by conventional plant breeding techniques.

CC

SO Sequence 787 BP; 209 A; 206 C; 216 G; 150 T; 6 other;

Query Match 72.2%; Score 16.6; DB 20; Length 787;

Best Local Similarity 82.6%; Pred. No. 82;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCAGACTGAGCAGCACCCTTT 23

DB 284 ACCTCGACTGGGAAAGCACCCTT 306

RESULT 5

AAV69440
 ID AAV69440 standard; cDNA: 809 BP.

AC AAV69440;

DT 15-MAR-1999 (first entry)

DE Banana fruit ripening-related clone U-7 cDNA.

KW Fruit ripening; banana; modulator; tissue senescence; crop; plant;

KM triploid; plant breeding; ss.

XX OS Musa acuminata.

XX PN WO9853085-A1.

XX PD 26-NOV-1998.

XX PF 05-MAY-1998; 98WO-GB01297.

XX PR 20-MAY-1997; 97GB-0010370.

XX PA (ZENE) ZENECA LTD.

XX PI Bird CR, Medina-Suarez RDU, Seymour GB;

XX DR WPI; 1999-059745/05.

XX PT New method of modulating fruit ripening or tissue senescence

XX PT characteristics of Musa plants - by introducing DNA sequences,

XX PT useful in processes for modifying plant/fruit ripening

XX PT characteristics

XX PS Claim 1; Page 18; 78pp; English.

XX CC AAV69440-V69512 are cDNA sequences which are used in a method to

XX CC modulate the fruit ripening or tissue senescence characteristics of Musa

XX CC acuminata (banana) plants. The method provides a recombinant way of

XX CC modulating ripening/senescence characteristics of bananas, which are a

XX CC globally important crop. In particular, the method can modulate such

XX CC characteristics in the desert banana, which due to its triploid nature

XX CC is unaffected by conventional plant breeding techniques.

XX SO Sequence 809 BP; 210 A; 214 C; 217 G; 153 T; 15 other;

Query Match 72.2%; Score 16.6; DB 20; Length 809;

Best Local Similarity 82.6%; Pred. No. 82;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCAGACTGAGCAGCACCCTTT 23

DB 277 ACCTCGACTGGGAAAGCACCCTT 299

RESULT 6

AAV27502

ID AAV27502 standard; cDNA: 1181 BP.

AC AAV27502;

DT 26-MAY-1999 (first entry)

DE Banana ethylene-forming enzyme cDNA sequence.

KW 1-aminocyclopropane-1-carboxylic acid synthase; ACS; EFE; banana;

KW ethylene forming enzyme; ethylene biosynthesis; plant; fruit ripening;

KW transgenic; enzyme; inhibition; flavour; texture; ss.

XX Musa sp.

XX US5886164-A.

XX PD 23-MAR-1999.

XX PF 15-APR-1996; 96US-0632598.

XX PR 15-APR-1996; 96US-0632598.

XX PA (ZENE) ZENECA LTD.

XX PI Bird CR, Fletcher JD;

XX DR WPI; 1999-228611/19.


```

FT      misc_signal      /label= sis responsive element
FT      162..168        /tag= j
FT      enhancer         /label= Ap1 binding site/TPA responsive element
FT      171..179        /tag= k
FT      /label= SV40 enhancer region Ap4
XX
XX      WO9510185-A.
XX
XX      20-APR-1995.
XX
XX      13-OCT-1994;    94WO-US11558.
XX
XX      15-OCT-1993;    93US-0136207.
XX      20-SEP-1994;    94US-0136207.
XX
XX      (UYAL-) UNIV ALABAMA.
XX      (UYDU-) UNIV DUKE.
XX
XX      Crapo JD, Day BJ, Foiz RJ, Freeman BA, Fridovich I;
XX      Oury T, Foltz J;
XX
XX      WPI, 1995-161483/21.
XX      P-PSDB; AAR72413.
XX
XX      Super:oxide dismutase mimetic(s) comprising a nitrogen-containing
XX      macrocyclic moiety - useful in treatment of inflammation,
XX      musculation disorders, atherosclerosis or gout.
XX
XX      Claim 33; Figure 24; 136pp; English.
XX
XX      Clone 7 from a human adult leukocyte genomic library was analysed.
XX      The analysis revealed that human EC-SOD contains three exons and
XX      two introns. The intron-exon boundaries are shown in AA088062-Q88065.
XX      AA088057 shows the entire sequence for the human EC-SOD gene. Exon 3
XX      contains the entire uninterrupted coding region. An obvious TATA
XX      box cannot be discerned for the EC-SOD gene. Two CAT box elements
XX      were identified. One is in the reverse orientation and located
XX      about 20bp upstream of the first exon. Transcriptional factor
XX      database searching of the 5' nontranslated region and first exon
XX      revealed several potential regulatory elements (see Fig.).
XX
XX      Sequence 10079 BP; 2482 A; 2613 C; 2407 G; 2577 T; 0 other;
XX
XX      Query Match      72.2%; Score 16.6; DB 16; Length 10079;
XX      Best Local Similarity 82.6%; Pred. No. 1.1e+02;
XX      Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX      QY      1 ACCACGACTGAGCAGCACCCTTT 23
XX      1238 ACCTCACTTAGCACCACCCTTT 1216
XX
XX      RESULT 11
XX      AAT92317/C
XX      ID      AAT92317 standard; cDNA; 10079 BP.
XX
XX      AC      AAT92317;
XX
XX      DT      06-FEB-1998 (first entry)
XX
XX      DE      Human EC-SOD gene.
XX
XX      KW      oxidant scavenger; extracellular matrix targeting moiety;
XX      cell surface matrix; nitrogen containing macrocyclic moiety;
XX      inflammatory condition; aberrant smooth muscle function;
XX      ischaemia reperfusion injury; myocardial infarction; stroke;
XX      acute head trauma; organ reperfusion; bowel ischaemia;
XX      pulmonary infarction; glaucoma; EC-SOD; superoxide dismutase;
XX      tetrameric glycosylated copper and zinc containing enzyme; ss.
XX      Homo sapiens.

```

```

XX      XX
XX      key      Location/Qualifiers
XX      CDS      5085..5807
XX      /tag= a
XX      /product= EC-SOD
XX
XX      WO9640223-A1.
XX
XX      19-DEC-1996.
XX
XX      07-JUN-1996;    96WO-US10497.
XX
XX      11-MAR-1996;    96US-0613418.
XX      07-JUN-1995;    95US-0476066.
XX
XX      (TROV/) TROVA M P.
XX      (UYAL-) UNIV ALABAMA.
XX      (UYDU-) UNIV DUKE.
XX
XX      Balnic-Haberle I, Crapo JD, Day BJ, Foiz RJ, Freeman BA;
XX      Fridovich I, Oury T, Trova MP;
XX
XX      WPI: 1997-077220/07.
XX      P-PSDB; AAM32599.
XX
XX      New porphyrin-type oxidant scavengers - used for protecting against
XX      oxidants and for modulating biological processes involving oxidants.
XX
XX      Claim 24; Fig 24; 195pp; English.
XX
XX      Oxidant scavengers have been developed comprising a nitrogen containing
XX      macrocyclic moiety and a cell surface or extracellular matrix targeting
XX      moiety, or their salts. The present sequence encodes the human EC-SOD,
XX      a tetrameric glycosylated copper and zinc containing superoxide
XX      dismutase enzyme. The oxidant scavengers can be used for protecting
XX      against the deleterious effects of oxidants and for modulating
XX      biological processes involving oxidants. They can be used for e.g.
XX      treating inflammatory conditions, treating disorders resulting from
XX      aberrant smooth muscle function or to protect against ischaemia
XX      reperfusion injuries associated with myocardial infarction, stroke,
XX      acute head trauma, organ reperfusion following transplantation, bowel
XX      ischaemia, pulmonary infarction, surgical occlusion of blood flow, and
XX      soft tissue injury. They can further be used to protect against damage
XX      to the eye due to sunlight (and to the skin) as well as glaucoma, and
XX      macular degeneration of the eye. Diseases of the bone are also amenable
XX      to treatment with the compounds, and connective tissue disorders
XX      associated with defects in collagen synthesis or degradation can be
XX      treated with the compounds. In the oxidant scavengers, substituents can
XX      be selected to render them resistant to degradation by haemoxigenase and
XX      also so that they do not interfere with normal porphyrin metabolism, can
XX      pass through cell membranes and bind to cell surface or extracellular
XX      matrix elements.
XX
XX      Sequence 10079 BP; 2481 A; 2612 C; 2409 G; 2577 T; 0 other;
XX
XX      Query Match      72.2%; Score 16.6; DB 18; Length 10079;
XX      Best Local Similarity 82.6%; Pred. No. 1.1e+02;
XX      Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX      QY      1 ACCACGACTGAGCAGCACCCTTT 23
XX      1238 ACCTCACTTAGCACCACCCTTT 1216
XX
XX      RESULT 12
XX      AAA28294/C
XX      ID      AAA28294 standard; DNA; 10079 BP.
XX
XX      AC      AAA28294;
XX
XX      DT      12-FEB-2001 (first entry)
XX
XX      DE      Human superoxide dismutase (EC-SOD) gene.

```

KX	Porphine; metal complex oxidant scavenger; inhibitor; EC-SOD; gout;
KW	superoxide dismutase; myocardial infarction; stroke; acute head trauma;
KM	organ reperfusion; bowel ischaemia; pulmonary infarction; glaucoma;
KM	skeletal muscle reperfusion injury; central nervous system disease; AIDS;
KM	dementia; stroke; amyotrophic lateral sclerosis; Parkinson's disease;
KM	Huntington's disease; neurological disorder; arthritis; hypertension;
KW	atherosclerosis; oedema; septic shock; pulmonary hypertension; asthma;
KW	impotence; infertility; endometriosis; diabetes; pneumonia; human;
KM	cystic fibrosis; sinusitis; autoimmune disease; ds.
OS	
XX	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	CDS
FT	5085..5807
FT	/tag= a
FT	/product= "EC-SOD"
FT	/note= "Superoxide dismutase"
FT	sig_peptide
FT	5085..5138
FT	/tag= b
FT	mat_peptide
FT	5139..5804
FT	/tag= c
XX	
PA	US6127356-A.
XX	
XX	03-OCT-2000.
PD	
PF	07-JUN-1996; 96US-0663028.
XX	
PR	15-OCT-1993; 93US-0136207.
PR	13-OCT-1994; 94US-0322766.
PR	07-JUN-1995; 95US-0476866.
PR	11-MAR-1996; 96US-0613418.
XX	
XX	(UYDU-) UNIV DUKE.
PI	
P1	Crapo JD, Fridovich I, Cury T, Foiz RJ, Trova MP, Freeman BA;
P1	Batlinic-Haberle I, Day BJ;
DR	WPI; 2000-664150/64.
DR	P-PSDB; AAY94782.
XX	
PT	New metal complexes of methine substituted porphines useful as
XX	catalytic oxygen scavengers -
XX	
PS	Disclosure; Fig 24; 97pp; English.
CC	This invention relates to porphines and their metal complex oxidant
CC	scavengers, where the metal is manganese, copper or iron. The porphines
CC	exhibit cardiant; cerebroprotective; vasotropic; ophthalmological;
CC	antiParkinsonian; nootropic; anticonvulsant; cyostatic; gynecological;
CC	antiarteriosclerotic; antiinflammatory; antibacterial; immunosuppressive;
CC	hypertensive; antidiabetic; antigout; antistamtic; and vitucide
CC	activity. The porphines act as superoxide dismutase (SOD) inhibitors and
CC	are used as catalytic scavengers of reactive oxygen species to protect
CC	against ischaemic reperfusion injuries associated with myocardial
CC	infarction, stroke, acute head trauma, organ reperfusion following
CC	transplantation, bowel ischaemia, pulmonary infarction, surgical
CC	occlusion of blood flow, soft tissue injury, skeletal muscle reperfusion
CC	injuries, glaucoma, macular degeneration of the eye, diseases of the
CC	bones, to increase the limited storage viability of transplanted hearts,
CC	kidneys, skin and other organs and tissues. The compounds are also useful
CC	in the treatment of diseases of the central nervous system (including
CC	AIDS dementia, stroke, amyotrophic lateral sclerosis), Parkinson's
CC	disease, Huntington's disease, disease of the musculature, cardiac
CC	fatigue of congestive heart failure, muscle weakness syndrome associated
CC	with myopathies, neurological disorders, arthritis, systemic
CC	hyperextension, atherosclerosis, oedema, septic shock, pulmonary
CC	contractions, microbial infections, gout, Type II diabetes mellitus,
CC	inflammation of the lungs, asthma, pneumonia, cystic fibrosis, chronic
CC	sinusitis and autoimmune disease. The present sequence represents the
CC	human EC-SOD gene. EC-SOD is a tetrameric glycosylated copper and zinc

CC containing enzyme. The enzyme is used to illustrate the activity of the
CC porphines of the invention.
XX

SQ Sequence 10079 BP; 2482 A; 2612 C; 2408 G; 2577 T; 0 other:

Query Match 72.2%; Score 16.6; DB 21; Length 10079;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Caps 0

Y 1 ACCAGACTGAGCAGCACTTT 23
111 111 1111 1111111

Db 1238 ACCTGACTGAGCACCCACCTT 1216

RESULT 13
AAV52205/c
ID AAV52205 standard; DNA: 14872 BP.

AC AAV52205;
XX
DT 23-OCT-1998 (first entry)

XX Streptococcus pneumoniae genome fragment SEQ ID NO:72.
DE
XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KM computer readable medium; vaccine; pharmaceutical composition; ds.
KW
XX Streptococcus pneumoniae.
CS
XX WO9818931-A2.
FN
PD 07-MAY-1998.
XX
PE 30-OCT-1997; 97WC-US19588.
XX
PR 31-OCT-1996; 96US-0029960.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
PI Kunsch CA, Rosen CA;
PI
PT pneumonae
PT pharmaceutical compositions and vaccines for Streptococcus
PT
DR WPI; 1998-272225/24.
XX
XX Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumonae
PS
XX

Claim 1: Page 593-602; 1409pp; English.

The present invention describes a computer readable medium which has
the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
recorded on it, or a representative fragment or a sequence at least 95%
identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in
SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
Streptococcus pneumoniae. The present invention also describes an
isolated nucleic acid molecule encoding a homologous of any of the
fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
nucleic acid molecule is produced by a process comprising: (a) screening
a genomic DNA library using as a probe a target sequence defined by any
of the sequences in SEQ ID NO:1 to 391, identifying members of the
library which contain sequences that hybridise to the target sequence and
isolating the nucleic acid molecules from the members; or (b) isolating
mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
molecules whose nucleotide sequence is homologous to amplification
primers derived from the fragment of the S. pneumoniae genome to prime
the amplification and isolating the amplified sequences. The computer
readable medium can be used in a computer-based system for identifying
fragments of the S. pneumoniae genome of commercial importance, or
expressions modulating fragments of the S. pneumoniae genome. Products
from the present invention can be used in diagnosis kits and assays, and
pharmaceutical compositions and vaccines for S. pneumoniae.

XX Sequence 14872 BP; 4340 A; 2874 C; 3210 G; 4448 T; 0 other;
 SO Query Match 72.2%; Score 16.6; DB 19; Length 14872;
 Best Local Similarity 82.6%; Pred. No. 1.1e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ACCACGACTGACGACGACCTT 23
 DB 11432 ACCAGGACTGACGACGACCTT 11410
 RESULT 14
 ID ABL27153 standard; DNA: 744 BP.
 AC ABL27153;
 XX ABL27153;
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32932.
 XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS
 XX MO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001MO-US09231.
 PE
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI: 2001-656860/75.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Claim 1; SEQ ID NO 32932; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 744 BP; 215 A; 205 C; 202 G; 122 T; 0 other;
 SQ
 Query Match 70.4%; Score 16.2; DB 23; Length 744;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ACCACGACTGACGACGACCT 21
 DB 682 ACCAAGACAGAGGACGACCT 702
 RESULT 15
 ABL27146

ID ABL27146 standard; DNA: 2510 BP.
 XX ABL27146;
 AC ABL27146;
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32911.
 XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS
 XX MO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001MO-US09231.
 PE
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI: 2001-656860/75.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Claim 1; SEQ ID NO 32911; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 2510 BP; 680 A; 581 C; 586 G; 663 T; 0 other;
 SQ
 Query Match 70.4%; Score 16.2; DB 23; Length 2510;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ACCACGACTGACGACGACCT 21
 DB 848 ACCAAGACAGAGGACGACCT 868
 Search completed: October 17, 2002, 08:50:51
 Job time : 160.222 secs

```

FT      misc_RNA      /product= 28S_rRNA
FT      5583..9396
FT      /*tag= f
FT      /function= spacer
FT      misc_RNA      7338..8291
FT      /*tag= g
FT      /function= probe
FT      /note= "claimed"
FT      misc_RNA      9397..9558
FT      /*tag= h
FT      /product= 18S_rRNA
XX
XX      JP07059577-A.
XX
XX      07-MAR-1995.
XX
XX      23-AUG-1993; 93JP-0227806.
XX
XX      23-AUG-1993; 93JP-0227806.
XX
XX      (KIRI ) KIRIN BEVERAGE KK.
XX
XX      WPI; 1995-135901/18.
XX
XX      Valencia orange ribosomal RNA gene - also probes and primers
XX      PT derived from citrus fruit RNA gene spacer regions, useful for
XX      PT RFLP analysis of citrus fruit
XX
XX      Claim 1; Page 7-10; 18pp; Japanese.
XX
XX      The ribosomal RNA gene has been isolated from Valencia oranges and
XX      CC sequenced (see AA088228). Primers based on sequences in the rRNA gene
XX      CC (see AA088228-Q088230 which are derived from nucleotides 1556-1575 and
XX      CC 2257-2276 of AA088228, respectively) can be used to amplify regions of
XX      CC rRNA genes from various different types of citrus fruits. RFLP
XX      CC analysis and DNA fingerprinting of the amplified fragments allows
XX      CC different species to be distinguished.
XX
XX      SO      Sequence 9558 BP; 1853 A; 2675 C; 2996 G; 2034 T; 0 other;
XX
XX      Query Match      80.9%; Score 17.8; DB 16; Length 9558;
XX      Best Local Similarity 90.5%; Pred. No. 20;
XX      Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY      1 CCGCTTCACTGCGCTTACTA 21
DB      2340 CCGTTCGCTGCGCTTACTA 2320

RESULT 15
AAF22281/c
ID      AAF22281 standard; DNA: 59590 BP.
XX
XX      AAF22281:
XX
XX      20-MAR-2001 (first entry)
XX
XX      DE      RAC containing repeats from centromeres 1-4 #4.
XX
XX      KM      Centromere; michrosome; vector; ds.
XX
XX      OS      Arabidopsis thaliana.
XX
XX      PN      W0200055325-A2.
XX
XX      PD      21-SEP-2000.
XX
XX      PF      17-MAR-2000; 2000WO-US07392.
XX
XX      PR      18-MAR-1999; 99US-0125219.
XX      01-APR-1999; 99US-0127409.
XX      18-MAY-1999; 99US-0134770.
XX      13-SEP-1999; 99US-0153584.

```

```

PR      17-SEP-1999; 99US-0154603.
XX
XX      (UYCH-) UNIV CHICAGO.
XX
XX      PI      Preuss D, Copenhagen G, Keith K;
XX
XX      DR      WPI; 2000-587529/55.
XX
XX      PT      Recombinant DNA construct comprising a plant centromere, useful for
XX      PT producing stably inherited michrosomes which can serve as vectors for
XX      PT the construction of transgenic plant and animal cells
XX
XX      PS      Claim 102; Page 351-364; 1449pp; English.
XX
XX      CC      The present invention relates to a recombinant DNA construct of a plant
XX      CC (Arabidopsis thaliana) centromere. The constructs are useful for
XX      CC producing stably inherited michrosomes which can serve as vectors for
XX      CC the construction of transgenic plant and animal cells expressing
XX      CC selected proteins such as hormones, enzymes, interleukins, clotting
XX      CC factors, cytokines, antibodies, and growth factors.
XX
XX      SO      Sequence 59590 BP; 17614 A; 12031 C; 13575 G; 16118 T; 252 other;
XX
XX      Query Match      80.9%; Score 17.8; DB 21; Length 59590;
XX      Best Local Similarity 90.5%; Pred. No. 23;
XX      Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY      1 CCGCTTCACTGCGCTTACTA 21
DB      45271 CCGTTCGCTGCGCTTACTA 45251

```

Search completed: October 17, 2002, 08:50:39
 Job time : 159.778 secs

```

XX Arabidopsis thaliana; rRNA gene; ribosomal DNA; intergenic region;
KW DNA construct; plant; multiple SalI repeat; stability; copy number;
KW transgene; antibody; insecticidal protein; Bt toxin; ss.
XX Synthetic.
XX WO9813505-A1.
XX 02-APR-1998.
XX 23-SEP-1997; 97WO-EP05217.
XX 24-SEP-1996; 96AT-0001695.
XX (PLB2 ) PLANT GENETIC SYSTEMS NV.
XX Bachmair A, Schweizer D;
XX WPI; 1998-286413/25.
XX DNA construct containing plant intergenic ribosomal DNA fragment
PI containing multiple SalI repeats; increases stability and copy
PT number of a transgene(s) in plants
XX Disclosure; Page 11; 48pp; English.
XX A DNA construct has been developed which comprises the following
CC operably linked DNA fragments: (a) a ribosomal DNA sequence, preferably
CC derived from a plant; (b) an expressible (especially plant-expressible)
CC promoter region; (c) a heterologous coding region; and (d) a
CC transcription termination and polyadenylation region which preferably is
CC active in plant cells. The present sequence represents an oligonucleotide
CC from the present invention. Transformed plant cells and transgenic plants
CC comprising the DNA constructs are used to produce a desired protein at a
CC high yield, e.g. antibodies, insecticidal proteins (e.g. a Bt toxin), a
CC protein useful in the food industry. Alternatively the constructs can be
CC used to produce an antisense RNA or ribozyme. The use of intergenic
CC ribosomal DNA enhances stability and the copy number or expression of
CC transgenes in a plant.
XX Sequence 24 BP; 3 A; 9 C; 6 G; 6 T; 0 other;
SO
Query Match 80.9%; Score 17.8; DB 19; Length 24;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CCGCTTCAGTCGCCGTACTA 21
DB 2 CCGCTTCAGTCGCCGTACTA 22

```

```

PR 20-MAY-1997; 97AU-0006903.
XX (INSE-) INSEARCH LTD.
XX Amoyal GS, Ellis JT, Rice CL;
XX WPI; 1999-045313/04.
XX Large subunit ribosomal DNA of Neospora species, especially N.
PT caninum - useful to derive nucleic acid sequences for isolation of
PT Neospora species by PCR amplification, e.g. to diagnose Neospora in
PT clinical samples
XX Claim 1; Fig 2; 40pp; English.
XX This nucleotide sequence encodes the large subunit (LSU) ribosomal
CC DNA (rDNA) of Neospora caninum. It is a compilation of overlapping
CC fragments obtained by PCR amplification of N. caninum (NC-Liverpool
CC strain) tachyzoite genomic DNA. N. caninum is a cyst-forming
CC coccidian parasite that causes neuromuscular disease in canines and
CC is a significant cause of abortion and neonatal mortality in
CC livestock. It is recognised as being closely related to Toxoplasma
CC gondii. Comparison of LSU rDNA of N. caninum with a consensus
CC sequence derived for the LSU rDNA of T. gondii demonstrated that
CC the D2 domain (C1/C1') region can serve as a target for the
CC development of a species-specific PCR (see AAV72869-75) for the
CC detection of Neospora rDNA, especially in clinical samples. In
CC addition, PCR MIMICS have been developed for use in competitive PCR
CC assays (see AAV72876-81) to differentially diagnose between Neospora
CC spp. and other related/non-related micro-organisms, particularly
CC Toxoplasma spp.
SO Sequence 3499 BP; 870 A; 734 C; 937 G; 958 T; 0 other;
Query Match 80.9%; Score 17.8; DB 20; Length 3499;
Best Local Similarity 90.5%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 CCGCTTCAGTCGCCGTACTA 21
DB 107 CCACTTACTCGCGCGTACTA 87

```

```

RESULT 13
AAV72868/C
ID AAV72868 standard; DNA; 3499 BP.
XX AAV72868;
XX 29-MAR-1999 (first entry)
XX Neospora caninum large subunit ribosomal DNA.
DE Neospora caninum large subunit ribosomal DNA.
XX large subunit ribosomal DNA; LSU-rDNA; detection; diagnosis; assay;
KW ss.
XX Neospora caninum (NC-Liverpool strain).
XX WO9853074-A1.
XX 26-NOV-1998.
XX 19-MAY-1998; 98WO-AU00367.

```

```

RESULT 14
AA088228/C
ID AA088228 standard; DNA; 9558 BP.
XX AA088228;
XX 24-NOV-1995 (first entry)
XX Valencia orange ribosomal RNA gene.
DE Valencia orange ribosomal RNA gene.
XX Ribosomal RNA; rRNA; 18S; 28S; 5.8S; Valencia orange; RFLP analysis;
KW DNA fingerprinting; restriction fragment length polymorphism; ds.
XX Citrus sp.
XX Key
XX misc_RNA
XX Location/Qualifiers
XX 1..1597
FT /*tag= a
FT /*product= 18S_rRNA
FT /*tag= b
FT /*tag= 1598..1848
FT /*function= spacer
FT 1849..2011
FT /*tag= c
FT /*product= 5.8S_rRNA
FT 2012..2238
FT /*tag= d
FT /*function= spacer
FT 2239..5582
FT /*tag= e

```

Db 222 CCGCTCAGTCGCCGCTACT 203

RESULT 10

AAV36211/C
ID AAV36211 standard; DNA: 4256 BP.

XX AAV36211:

DT 03-SEP-1998 (first entry)

DE Total contiguous sequence of P. carinii in immunosuppressed rats.

XX Detection; diagnosis: 26S rRNA gene; P. carinii specific; infection;

KM Species identification; ss.

OS pneumocystis carinii.

PN US5776680-A.

PD 07-JUL-1998.

PP 21-JUL-1995; 95US-0505509.

PR 30-JUL-1992; 92US-0922987.

PR 31-AUG-1994; 94US-0298087.

PA (UYNE-) UNIV NEW JERSEY.

PI Leibowitz MJ, Liu Y;

DR WPI: 1998-398016/34.

PT Detection of Pneumocystis carinii - by amplification of nucleic acid

PT from sample with PCR primers specific for the 26S rRNA gene of

PT Pneumocystis carinii

XX Disclosure: Columns 29-34: 42pp; English.

XX The present sequence represents the total contiguous sequence of

CC Pneumocystis carinii in immunosuppressed Sprague-Dawley rats. The

CC specification describes a method for the diagnosis of Pneumocystis

CC carinii which comprises detecting the presence of a nucleic acid

CC sequence containing the 26S rRNA gene specific for P. carinii in a

CC sample. The 26S rRNA gene in a sample is amplified, and the primer

CC extension products detected by hybridisation with a labelled

CC oligonucleotide. The methods can be used for the diagnosis of

CC P. carinii infection and for the detection of various species of

CC P. carinii.

SO Sequence 4256 BP; 1190 A; 766 C; 1124 G; 1176 T; 0 other;

Query Match 83.6%; Score 18.4; DR 19; Length 4256;

Best Local Similarity 95.0%; Pred. No. 9.2;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCGCTCAGTCGCCGCTACT 20

DB 660 CCACTTCACTCGCGCTACT 641

RESULT 11

AAV83003/C

ID AAV83003 standard; DNA: 4256 BP.

XX AAV83003:

XX 23-FEB-1999 (first entry)

XX Contiguous sequence determined for P. carinii from immunosuppressed rats.

DE PCR amplification; sequencing; assay; inhibitor; nuclear RNA gene;

XX

KW

KM catalytic Group I self-splicing intron reaction; drug screening; ss.

OS Pneumocystis carinii.

XX Key

FT Intron

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

FT exon

Location/Qualifiers

1..22

/tag= a

/note= "3'-terminal portion of intron 1"

23..53

/tag= b

/note= "exon 2 of 16S rRNA"

54..216

/tag= c

/note= "Internal transcribed spacer 1"

217..374

/tag= d

/note= "5.8S rRNA"

375..556

/tag= e

/note= "Internal transcribed spacer 2"

557..4256

/tag= f

/note= "26S rRNA"

US5849484-A.

15-DEC-1998.

19-JUN-1995; 95US-0491690.

27-MAY-1993; 93US-0068248.

30-JUL-1992; 92US-0922987.

19-JUN-1995; 95US-0491690.

(UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

Leibowitz MJ, Liu Y;

WPI: 1999-069716/06.

Screening assays for drugs against Pneumocystis carinii - based on

inhibition of 26S rRNA gene intron self-splicing

Example 1; Fig 2; 51pp; English.

XX The present sequence represents the contiguous sequence determined for

CC Pneumocystis carinii from immunosuppressed Sprague-Dawley rats, using

CC the in vitro method of the invention. The method assays for an inhibitor

CC of the catalytic Group I self-splicing intron reaction in the nuclear

CC RNA genes of P. carinii. The method is useful for screening potential

CC drugs for treating P. carinii infections before more costly animal

CC testing is conducted.

SO Sequence 4256 BP; 1190 A; 766 C; 1124 G; 1176 T; 0 other;

Query Match 83.6%; Score 18.4; DB 20; Length 4256;

Best Local Similarity 95.0%; Pred. No. 9.2;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCGCTCAGTCGCCGCTACT 20

DB 660 CCACTTCACTCGCGCTACT 641

RESULT 12

AAV31383

ID AAV31383 standard; DNA: 24 BP.

XX AAV31383:

XX 07-SEP-1998 (first entry)

XX Oligonucleotide Q SEQ ID NO:5 from W09813505.

DE

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC carcinoma, ductal carcinoma in situ or lobular carcinoma in situ. This
 CC sequence represents DNA encoding a human HER polypeptide.
 XX

SQ Sequence 3040 BP; 851 A; 686 C; 976 G; 527 T; 0 other;

Query Match 83.6%; Score 18.4; DB 22; Length 3040;
 Best Local Similarity 95.0%; Pred. No. 8.9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACT 20
 |||||||||||||
 DB 2780 CCGCTTCACTCGCCGTACT 2799

RESULT 8
 AA57868
 ID AA57868 standard; cDNA; 3073 BP.
 XX
 AC AA57868;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Protein regulating gene expression PRGE-30 cDNA clone 1977214.
 XX
 KW Protein regulating gene expression; PRGE-30; human;
 KW cell proliferation; antiproliferative; inflammation;
 KW antiinflammatory; gene therapy; diagnosis; RNA binding protein;
 KW RNP-1; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 122..2875
 FT /*tag= a
 XX
 PN MO9964596-A2.
 PD 16-DEC-1999.
 XX
 PF 11-JUN-1999; 99WO-US13281.
 XX
 PR 12-JUN-1998; 98US-0089029.
 PR 29-JUL-1998; 98US-0094575.
 PR 14-OCT-1998; 98US-0104624.
 XX
 PA (INCYTE-) INCYTE PHARM INC.
 PI Lal P, Yue H, Tang YT, Hillman Jr., Bandman O, Corley NC;
 PI Guegler KJ, Gorgone GA, Baughn MR, Patterson C, Lu DM;
 DR P-PSDB; AAY58637.
 XX
 PS WPI: 2000-116543/10.
 PT New human polypeptides that regulate gene expression, for treatment,
 PT prevention and diagnosis of, e.g. cancer -
 XX
 PS Claim 9; Page 146-147; 150pp; English.
 XX
 CC The present sequence is that of Incyte clone 1977214 encoding new
 CC human protein regulating gene expression PRGE-30 (see AAY58637). The
 CC cDNA was initially isolated from pancreatic tumour cDNA library
 CC PAKMT002, and the full-length sequence assembled from overlapping
 CC clones from a number of libraries. PRGE-30 is expressed in nervous,
 CC reproductive and gastrointestinal tissues associated with cell
 CC proliferation and inflammation diseases, disorders or conditions.
 CC It is characterised as an RNA binding protein (RNP-1). The invention
 CC provides PRGE polypeptides (see AAY58638) and polynucleotides (see
 CC AA57839-69), expression vectors, host cells, antibodies, agonists and
 CC antagonists. It also provides methods for diagnosing, treating or
 CC preventing disorders associated with expression of PRGE.
 CC Polynucleotides are also used as sources of probes and primers for
 CC diagnosis and monitoring of disease, also for detecting related
 CC sequences and in gene mapping.

XX
 SQ Sequence 3073 BP; 840 A; 699 C; 998 G; 536 T; 0 other;

Query Match 83.6%; Score 18.4; DB 21; Length 3073;
 Best Local Similarity 95.0%; Pred. No. 9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACT 20
 |||||||||||||
 DB 2854 CCGCTTCACTCGCCGTACT 2873

RESULT 9
 AA192675/C
 ID AA192675 standard; cDNA; 3130 BP.
 XX
 AC AA192675;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 12735.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 DR WPI: 2001-514838/56.
 DR P-PSDB; AAO12744.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 1; SEQ ID NO 12735; 1399pp + Sequence listing; English.
 XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity eluting to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3130 BP; 566 A; 1005 C; 703 G; 856 T; 0 other;

Query Match 83.6%; Score 18.4; DB 22; Length 3130;
 Best Local Similarity 95.0%; Pred. No. 9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACT 20

XX DR WPI: 1994-167493/20.
 XX CC Probe for detection of infectious disease - comprises DNA
 PT fragment specific for fungal disease agent, for diagnosis of e.g.
 PT Candida infection
 XX PS Claim 5; Page 19; 35pp; Japanese.
 CC CC Genomic DNA was isolated from *Candida albicans* (CA-26) (AA063962) and
 CC fragments were generated by an *Hpa*I digest. This is used as a probe
 CC for detection of *Candida* infection. The advantage of this probe
 CC is that diagnosis time can be reduced from 3-4 days to 1-2, while
 CC maintaining accuracy.
 XX CC
 SQ Sequence 899 BP; 233 A; 200 C; 224 G; 242 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 15; Length 899;
 Best Local Similarity 95.0%; Pred. No. 8.1;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 0Y 1 CCGCTTCACTGCGCGTACT 20
 Db 757 CCGCTTCACTGCGCGTACT 738
 RESULT 6
 AAX78193
 ID AAX78193 standard; cDNA; 3040 BP.
 AC AAX78193;
 XX 20-AUG-1999 (first entry)
 DT Human HET cDNA.
 DE
 XX HET; cytosolic; hsp27 promoter; suppressor; hsp27; tumour suppressor;
 KW hsp27-ERF-TAF-binding protein; diagnosis; prognosis; breast cancer;
 KW infiltrating duct carcinoma; lobular carcinoma; medullary carcinoma;
 KW mucinous carcinoma; tubular carcinoma; ductal carcinoma; therapy;
 KW screening; human; ss.
 XX
 OS Homo sapiens.
 XX W09928466-A1.
 PN 10-JUN-1999.
 PD 30-NOV-1998; 98WO-US25381.
 PF 04-DEC-1997; 97US-0068132.
 PR (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Allred CD, Fugua SAW, Lee AV, O'Connell P, Oesterreich S;
 PI Osborne CK;
 XX WPI: 1999-385382/32.
 DR P-PSDB: AAU08991.
 XX New isolated tumor suppressor, HET
 PT
 XX Claim 3; Page 126-127; 137pp; English.
 CC This invention describes a nucleic acid segment that encodes a novel
 CC human HET polypeptide which has cytosolic activity. The HET protein
 CC binds to the hsp27 promoter in order to suppress expression of the
 CC hsp27 protein. The HET protein (hsp27-ERF-TAF-binding protein) can bind
 CC to the hsp27 promoter and suppress expression of hsp27. The HET acts as a
 CC tumour suppressor gene. The products and methods can be used for the
 CC diagnosis and prognosis of breast cancers including infiltrating duct
 CC carcinoma, lobular carcinoma, medullary carcinoma, mucinous carcinoma,
 CC tubular carcinoma, ductal carcinoma in situ and lobular carcinoma in

CC situ. The products can also be used for therapy and in the screening of
 CC compounds for activity in either stimulating HET activity, overcoming the
 CC lack of HET or blocking the effect of a mutant HET molecule.
 XX CC
 SQ Sequence 3040 BP; 851 A; 686 C; 976 G; 527 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 20; Length 3040;
 Best Local Similarity 95.0%; Pred. No. 8.9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 0Y 1 CCGCTTCACTGCGCGTACT 20
 Db 2780 CCGCTTCACTGCGCGTACT 2799
 RESULT 7
 AAS12099
 ID AAS12099 standard; DNA; 3040 BP.
 AC AAS12099;
 XX 21-NOV-2001 (first entry)
 DT Human HET DNA sequence.
 DE
 XX HET; nuclear matrix protein; antioestrogen resistance; breast cancer; ds;
 KW scaffold attachment factor B; SAF-B; loss of heterozygosity; human; yhm;
 KW infiltrating duct carcinoma; lobular carcinoma; medullary carcinoma;
 KW mucinous carcinoma; tubular carcinoma; ductal carcinoma;
 KW chromosome 19p13.2-p13.3.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 54..2801
 FT /*tag= a
 FT /product= "Human HET protein"
 PN W0200163292-A2.
 XX 30-AUG-2001.
 PD 22-FEB-2001; 2001WO-US06135.
 PF 22-FEB-2000; 2000US-0184097.
 PR (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Oesterreich S, Osborne CK, Lee AV, Fugua SA;
 PI Osborne CK;
 XX WPI: 2001-536660/59.
 DR P-PSDB: AAU07605.
 XX Detection of antioestrogens breast cancer cells comprises the
 PT measurement of a HET polypeptide -
 PT
 XX Examples: Page 137; 140pp; English.
 CC The invention relates to the detection of resistance to antioestrogens in
 CC breast cancer cells, comprising obtaining a breast cancer cell sample,
 CC contacting the sample with an antibody that specifically binds to the
 CC nuclear matrix protein, HET, (also known as scaffold attachment factor B
 CC or SAF-B) to form a complex and measuring the amount of HET polypeptide
 CC present. A decrease in HET level in the sample compared to the control
 CC indicates a presence of antioestrogen resistance. Resistance in breast
 CC cancer cells can be predicted by comparing the amount of HET protein in
 CC samples from patients with antioestrogen-resistant and
 CC antioestrogen-sensitive breast cancers. Malignant breast cancers can be
 CC diagnosed by determining loss of heterozygosity (LOH) at a chromosome
 CC locus where the LOH at the locus is indicative of antioestrogen
 CC resistance in breast cancer cells, as well as a decrease in chance of
 CC survival. The cancerous cells may be from an infiltrating duct carcinoma,
 CC lobular carcinoma, medullary carcinoma, mucinous carcinoma, tubular

PR 24-NOV-1987; 87WO-US03009.
 PR 11-DEC-1991; 91US-0806929.
 PR 24-NOV-1986; 86US-0934244.
 PR 07-AUG-1987; 87US-0083542.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI McDonough SH, Kop JA, Smth RD, Hogan JJ;
 DR WPI: 2001-060029/07.
 XX
 PI Preparing a probe for nucleic acid hybridization assays comprises
 PT constructing a nucleotide polymer sufficiently complementary to
 PT hybridize to an rRNA region that distinguishes non-viral target from
 PT non-viral non-target species -
 XX
 PS Disclosure; Fig 5; 75pp; English.
 CC
 CC The present invention provides novel methods of producing probes for use
 CC in the identification of a number of microorganisms. These include E.
 CC coli, Mycobacteria, Mycoplasma, Campylobacter, Chlamydia, Enterobacter,
 CC Legionella, Salmonella, Pseudomonas, Neisseria gonorrhoeae, fungi and
 CC bacteria.
 CC
 SO Sequence 3550 BP; 941 A; 697 C; 1000 G; 912 U; 0 other;
 Query Match 95.5%; Score 21; DB 22; Length 3550;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGCTTCACTGCGCGTTACTA 21 22
 DB 261 CCGCTTCACTGCGCGTTACTA 241
 RESULT 4
 AAQ39050/C
 ID AAQ39050 standard; DNA: 6824 BP.
 XX
 AC AAQ39050:
 XX
 DT 28-JUL-1993 (first entry)
 XX
 DE K.lactis/S. cerevisiae genetic vector.
 XX
 KM Genetic; vector: integration; Kluyveromyces lactis; 255 ribosomal DNA;
 KM Saccharomyces cerevisiae; E. coli; domain; yeast; plasmid; promoter;
 KM expression cassette; HIS3; marker; transformant; human; lysozyme; HLZ;
 KM GAL7; signal sequence; killer toxin; transcription termination signal;
 KM FLP; 2 micron plasmid; ss.
 XX
 OS Synthetic.
 XX
 PN EP537456-A.
 XX
 PD 21-APR-1993.
 XX
 PF 31-AUG-1992; 92EP-0114838.
 XX
 PR 04-SEP-1991; 91IT-0M12349.
 XX
 PA (ISTS) SCLAVO SPA.
 XX
 PI Galeotti CL, Gallo E, Riccio ML, Rossolini GM, Thaller MC;
 DR WPI: 1993-127394/16.
 XX
 PT Vector for Kluyveromyces lactis and Saccharomyces cerevisiae -
 PT which allows stable multiple integration of DNA for prodn. of
 PT heterologous proteins
 XX
 PS Claim 1; Fig 1; 26pp; English.
 XX

CC This sequence represents a genetic vector which allows the stable
 CC multiple integration of DNA sequences into the genome of Kluyveromyces
 CC lactis and Saccharomyces cerevisiae. This sequence can be used in an
 CC integrating vector which comprises a region necessary for the stable
 CC maintenance of the plasmid in E. coli and a domain which acts as an
 CC integrating unit consisting of two not contiguous sequences of the 255
 CC ribosomal DNA from S. cerevisiae, flanking a genetic marker suitable
 CC for selection of the yeast transformants in which the integration
 CC event has occurred. Other DNA sequences may be introduced into the
 CC integration plasmid, such as expression cassettes. The gene HIS3
 CC from K. lactis and S. cerevisiae is prefl. used as a genetic marker
 CC for the selection of transformants and an expression cassette for the
 CC production and secretion into the culture medium of human lysozyme.
 CC This complete transformation vector is 7850 bp long and includes the
 CC integration vector of the invention and an expression cassette
 CC comprising the K. lactis GAL7 promoter, the signal sequence of the K.
 CC lactis killer toxin, the cDNA encoding the ripe form of human lysozyme
 CC (HLZ) and the transcription termination signal FLP of the 2 micron
 CC plasmid from S. cerevisiae.
 CC
 SO Sequence 6824 BP; 1815 A; 1521 C; 1726 G; 1762 T; 0 other;
 Query Match 95.5%; Score 21; DB 14; Length 6824;
 Best Local Similarity 100.0%; Pred. No. 0.48;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGCTTCACTGCGCGTTACTA 21
 DB 144 CCGCTTCACTGCGCGTTACTA 124
 RESULT 5
 AAQ63962/C
 ID AAQ63962 standard; DNA: 699 BP.
 XX
 AC AAQ63962:
 XX
 DT 30-NOV-1994 (first entry)
 XX
 DE Candida albicans (CA-26) genomic DNA fragment.
 XX
 KM Detection of infectious disease; fungal disease agent;
 KM probes; diagnosis; ds.
 XX
 OS Candida albicans.
 XX
 FH Key Location/Qualifiers
 FT misc-feature 14..202
 FT /*tag- a
 FT /note- "fragment isolated and used as a template
 FT for probe formation"
 FT misc-feature 372..737
 FT /*tag- b
 FT /note- "fragment isolated and used as a template
 FT for probe formation"
 FT misc-feature 517..737
 FT /*tag- c
 FT /note- "fragment isolated and used as a template
 FT for probe formation"
 XX
 PN WO9410341-A.
 XX
 PD 11-MAY-1994.
 XX
 PF 25-OCT-1993; 93WO-JP01555.
 XX
 PR 23-OCT-1992; 92JP-0285802.
 XX
 PA (FUSO) FUSO YAKURIN KOGYO KK.
 PA (ONOF) ONO Y.
 PA (FUSO) FUSO PHARM IND LTD.
 XX
 PI Hirotsu T, Keshi H, Matsuhisa A, Ohno T;

XX Claim 1a: 49; 130pp; German.

XX This invention describes novel cDNA molecules corresponding to tobacco
CC (Nicotiana tabacum) genes that are expressed in response to tobacco
CC mosaic virus (TMV) infection. Regulatory regions that provide specific
CC expression of the nucleic acids in connection with induced resistance
CC can be used to identify compounds useful for plant protection by
CC transforming a cell with a recombinant DNA molecule containing a
CC selectable and/or detectable marker linked to the regulatory region,
CC culturing the cell in the presence of a test compound or sample, and
CC identifying a compound or sample that suppresses or activates and/or
CC enhances expression of the marker. Host cells containing the nucleic
CC acid, or polypeptides encoded by the nucleic acids, can be used to
CC identify compounds that inhibit or activate a polypeptide involved in
CC induced resistance, by contacting the cell or polypeptide with one or
CC more test compounds and identifying compound(s) that specifically
CC modulate the activity of the polypeptide. Plant protection agents can be
CC prepared by formulating the identified compounds in a form suitable for
CC their uptake and optionally metabolism in organisms, especially
CC plants, phytopathogenic fungi, nematodes or insect pests. The nucleic
CC acids or regulatory regions, or corresponding vectors, or the
CC polypeptides encoded by the nucleic acids, or compounds identified as
CC above, can be used to improve the pathogen resistance of plants or to
CC produce transgenic plants. The nucleic acids or regulatory regions, or
CC vectors, cells, plants or plant tissues containing them, or the
CC polypeptides and antibodies to the polypeptides, can be used to screen
CC for new plant protection agents or genes involved in induction of
CC resistance in plants. Vectors containing the nucleic acids, or the
CC regulatory regions or recombinant DNA molecule or vectors containing
CC them, or the compounds identified as above, or the polypeptides or
CC antibodies, can be used in diagnostic compositions. AA233676-233688
CC represent fragments of tobacco cDNA which are associated with plant
CC resistance.

XX Sequence 218 BP; 68 A; 43 C; 65 G; 42 T; 0 other;

XX Query Match 100.0%; Score 22; DB 20; Length 218;

XX Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Gaps 0;

XX Matches 22; Conservative 0; Indels 0; Gaps 0;

0Y 1 CCGCTTCATCGCGCTTACTAG 22
|||||
DB 59 CCGCTTCATCGCGCTTACTAG 38

RESULT 2

AAf09568/c

ID AAF09568 standard; cDNA; 431 BP.

XX AAF09568;

DT 13-MAR-2001 (first entry)

DE Fusarium venenatum EST SEQ ID No. 2091.

XX Multiple gene expression; filamentous fungal cell; EST;

XX expressed sequence tag; Fusarium venenatum; Aspergillus niger;

XX Aspergillus oryzae; Trichoderma reesei; identification; recombination;

XX culture condition; environmental stress; spore morphogenesis;

XX metabolic pathway engineering; catabolic pathway engineering; ss.

XX Fusarium venenatum.

XX WO200056762-A2.

XX 28-SEP-2000.

XX 22-MAR-2000; 2000MO-US07781.

XX 22-MAR-1999; 99US-0273623.

XX (NOVO) NOVO NORDISK BIOTECH INC.

PA (NOVO) NOVO NORDISK AS.

XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IS, Olsen PB;

XX WPI: 2000-594572/56.

P7 Monitoring differential expression of genes in filamentous fungal cells
P7 uses fluorescence-labeled nucleic acids isolated from the cells and a
P7 substrate of expressed sequence tags -

XX Claim 86: Page 1160; 3161pp; English.

CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organization of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.

XX Sequence 431 BP; 139 A; 78 C; 115 G; 96 T; 3 other;

XX Query Match 100.0%; Score 22; DB 21; Length 431;

XX Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Gaps 0;

XX Matches 22; Conservative 0; Indels 0; Gaps 0;

0Y 1 CCGCTTCATCGCGCTTACTAG 22
|||||
DB 85 CCGCTTCATCGCGCTTACTAG 64

RESULT 3

AAf23019/c

ID AAF23019 standard; rRNA; 3550 BP.

XX AAF23019;

DT 20-MAR-2001 (first entry)

DE Yeast 28S rRNA sequence.

XX Probe: PCR primer; 5S rRNA; 16S rRNA; 23S rRNA; 28S rRNA; 18S rRNA;

XX Mycobacterium; Enteroococcus; Chlamydia; Mycoplasma; E. coli; Legionella;

XX Salmonella; Pseudomonas; Campylobacter; Neisseria gonorrhoeae; fungus;

XX bacterium; ss.

XX Saccharomyces cerevisiae.

XX US6150517-A.

XX 21-NOV-2000.

XX 30-MAR-1995; 95US-0454063.

XX 22-FEB-1994; 94US-0200866.

XX 24-NOV-1987; 87US-0295208.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 07:07:41 : Search time 31.2889 Seconds

(without alignments)
172.711 Million cell updates/sec

Title: US-09-780-113d-8

Perfect score: 22

Sequence: 1 ccgcctactcgcgcctactag 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 12281652 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	95.5	666	4 US-08-998-416-1013	Sequence 1013, App
C 2	21	95.5	697	4 US-08-998-416-909	Sequence 909, App
C 3	21	95.5	699	4 US-08-998-416-774	Sequence 774, App
C 4	21	95.5	700	4 US-08-998-416-813	Sequence 813, App
C 5	21	95.5	720	4 US-08-998-416-601	Sequence 601, App
C 6	18.4	83.6	899	1 US-08-416-831B-1	Sequence 1, Appl
C 7	18.4	83.6	4256	2 US-08-505-509-31	Sequence 31, Appl
C 8	18.4	83.6	4256	2 US-08-491-690A-31	Sequence 31, Appl
C 9	16.8	76.4	1703	4 US-08-378-313-18	Sequence 18, Appl
C 10	16.8	76.4	9060	4 US-08-378-313-20	Sequence 20, Appl
C 11	16	72.7	33	1 US-08-308-461-3	Sequence 3, Appl
C 12	16	72.7	33	5 PCT-US95-11823-3	Sequence 3, Appl
C 13	15.4	70.0	1627	3 US-09-106-464-1	Sequence 1, Appl
C 14	15	68.2	20	2 US-08-778-912A-9	Sequence 9, Appl
C 15	15	68.2	20	2 US-08-778-912A-11	Sequence 11, Appl
C 16	15	68.2	20	2 US-08-778-912A-15	Sequence 15, Appl
C 17	15	68.2	20	4 US-09-541-941B-9	Sequence 9, Appl
C 18	15	68.2	20	4 US-09-541-941B-11	Sequence 11, Appl
C 19	15	68.2	20	4 US-09-541-941B-15	Sequence 15, Appl
C 20	15	68.2	815	4 US-09-541-941B-20	Sequence 20, Appl
C 21	15	68.2	870	2 US-08-778-912A-4	Sequence 4, Appl
C 22	15	68.2	874	2 US-09-541-941B-4	Sequence 4, Appl
C 23	15	68.2	874	2 US-08-778-912A-1	Sequence 1, Appl
C 24	15	68.2	874	4 US-09-541-941B-2	Sequence 2, Appl
C 25	15	68.2	875	2 US-08-778-912A-2	Sequence 2, Appl
C 26	15	68.2	875	2 US-08-778-912A-3	Sequence 3, Appl
C 27	15	68.2	875	2 US-08-778-912A-5	Sequence 5, Appl

28	15	68.2	875	2 US-08-778-912A-7	Sequence 7, Appl
29	15	68.2	875	4 US-09-541-941B-1	Sequence 1, Appl
30	15	68.2	875	4 US-09-541-941B-3	Sequence 3, Appl
31	15	68.2	875	4 US-09-541-941B-5	Sequence 5, Appl
32	15	68.2	875	4 US-09-541-941B-7	Sequence 7, Appl
33	15	68.2	876	2 US-08-778-912A-6	Sequence 6, Appl
34	15	68.2	876	4 US-09-541-941B-6	Sequence 6, Appl
35	15	68.2	898	4 US-09-541-941B-22	Sequence 22, Appl
36	15	68.2	898	4 US-09-541-941B-23	Sequence 23, Appl
37	15	68.2	898	4 US-09-541-941B-24	Sequence 24, Appl
C 38	15	68.2	898	4 US-09-541-941B-25	Sequence 25, Appl
C 39	15	68.2	904	4 US-09-541-941B-21	Sequence 21, Appl
C 40	15	68.2	917	4 US-09-541-941B-18	Sequence 18, Appl
C 41	15	68.2	918	4 US-09-541-941B-16	Sequence 16, Appl
C 42	15	68.2	918	4 US-09-541-941B-17	Sequence 17, Appl
C 43	15	68.2	918	4 US-09-541-941B-19	Sequence 19, Appl
C 44	14.6	66.4	681	1 US-08-378-588-20	Sequence 20, Appl
C 45	14.6	66.4	681	2 US-08-811-094-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-998-416-1013/C
Sequence 1013, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Red: schung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESSES:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC 27709
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CCG1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8689
TELEFAX: 919-541-8687
INFORMATION FOR SEQ ID NO: 1013:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PA61621UP

US-08-998-416-1013

Query Match 95.5%; Score 21; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACTA 21
|||||
DB 457 CCGCTTCACTCGCCGTACTA 437

RESULT 2

US-08-998-416-909
; Sequence 909, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 909:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1567RP
; US-08-998-416-909

Query Match 95.5%; Score 21; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACTA 21
|||||
DB 332 CCGCTTCACTCGCCGTACTA 352

RESULT 3

US-08-998-416-774/c

; Sequence 774, Application US/08998416
; Patent No. 6239264

; GENERAL INFORMATION:

; APPLICANT: Philippsen, Peter
; APPLICANT: Pohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 774:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1498UP
; US-08-998-416-774

Query Match 95.5%; Score 21; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACTA 21
|||||
DB 457 CCGCTTCACTCGCCGTACTA 437

RESULT 4

US-08-998-416-813
; Sequence 813, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII

TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 813:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1518RP
US-08-998-416-813

Query Match 95.5%; Score 21; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTA 21
|||||
Db 332 CCGCTTCACTCGCGCTTACTA 352

RESULT 5
US-08-998-416-601
Sequence 601 Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,831B
FILING DATE: 19-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/01555
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPA 4-285802
FILING DATE: 23-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 19036/32578
TELECOMMUNICATION INFORMATION:

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 601:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1411RP
US-08-998-416-601

Query Match 95.5%; Score 21; DB 4; Length 720;
Best Local Similarity 100.0%; Pred. No. 0.058;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTA 21
|||||
Db 332 CCGCTTCACTCGCGCTTACTA 352

RESULT 6
US-08-416-831B-1/c
Sequence 1, Application US/08416831B
Patent No. 5708159
GENERAL INFORMATION:
APPLICANT: Hirotsu, Tsuneya
APPLICANT: Keshi, Hiroyuki
APPLICANT: Matsuhisa, Akio
TITLE OF INVENTION: Probe for Diagnosing Infectious Diseases
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,831B
FILING DATE: 19-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/01555
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPA 4-285802
FILING DATE: 23-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 19036/32578
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 899 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Candida albicans
STRAIN: Clinical isolate CA-26
US-08-416-831B-1

Query Match 83.6%; Score 18.4; DB 1; Length 899;
Best Local Similarity 95.0%; Pred. No. 1.2;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCCCGCTTACT 20
|||||
Db 757 CCGCTTCACTCCCGCTTACT 738

RESULT 7
US-08-505-509-31/C
Sequence 31, Application US/08505509
Patent No. 5776680
GENERAL INFORMATION:
APPLICANT: Liebowitz, Michael J.
APPLICANT: Liu, Yong
TITLE OF INVENTION: Diagnostic Probes for
TITLE OF INVENTION: Pneumocystis Carinii
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard R. Muccino
STREET: P.O. Box 1267
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08551
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,509
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/298,087
FILING DATE:
APPLICATION NUMBER: US/07/922,987
FILING DATE: 30-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD1-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 466-3407
TELEFAX: (609) 466-2760
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 4256 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-505-509-31

Query Match 83.6%; Score 18.4; DB 1; Length 4256;
Best Local Similarity 95.0%; Pred. No. 1.4;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCGCTTCACTCCCGCTTACT 20
|||||
Db 660 CCGCTTCACTCCCGCTTACT 641

RESULT 8
US-08-491-690A-31/C
Sequence 31, Application US/08491690A
Patent No. 5849484
GENERAL INFORMATION:
APPLICANT: Liebowitz, Michael J.
APPLICANT: Liu, Yong
TITLE OF INVENTION: In Vitro Assay For Inhibitors
TITLE OF INVENTION: Of the Intron Self-Splicing Reaction in Pneumocystis Carinii
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard R. Muccino
STREET: 758 Springfield Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,690A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,248
FILING DATE: 27-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD1-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 273-4988
TELEFAX: (908) 273-4679
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 4256 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-491-690A-31

Query Match 83.6%; Score 18.4; DB 2; Length 4256;
Best Local Similarity 95.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCCCGCTTACT 20
|||||
Db 660 CCGCTTCACTCCCGCTTACT 641

RESULT 9
US-08-378-313-18
Sequence 18, Application US/08378313
Patent No. 6207881
GENERAL INFORMATION:
APPLICANT: THEOLOGIS, ATHANASIOS
APPLICANT: SATO, TAKAHIDO
TITLE OF INVENTION: CONTROL OF FRUIT RIPENING THROUGH
TITLE OF INVENTION: GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,313
FILING DATE:
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/862,493
FILING DATE: 02-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 29190-20002.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 856-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1703 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 11..1489
US-08-378-313-18

Query Match 76.4%; Score 16.8; DB 4; Length 1703;
Best Local Similarity 90.0%; Pred. No. 8.1;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CGCTTCACGCGCGTTACTA 21
Db 1458 CGCCTCAGCGCGTTACTA 1477

RESULT 10
US-08-378-313-20
Sequence 20, Application US/08378313
Patent No. 6207881
GENERAL INFORMATION:
APPLICANT: THEOLOGIS, ATHANASIOS
APPLICANT: SATO, TAKAHIDO
TITLE OF INVENTION: CONTROL OF FRUIT RIPENING THROUGH
NUMBER OF SEQUENCES: 34
TITLE OF INVENTION: GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,313
FILING DATE:
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/862,493

FILING DATE: 02-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 29190-20002.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 856-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 9060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: join(2704..2880, 2968..3099, 3183..3344, 3810
LOCATION: ..4376, 4463..4903)
US-08-378-313-20

Query Match 76.4%; Score 16.8; DB 4; Length 9060;
Best Local Similarity 90.0%; Pred. No. 9.7;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CGCTTCACGCGCGTTACTA 21
Db 4872 CGCCTCAGCGCGTTACTA 4891

RESULT 11
US-08-308-461-3
Sequence 3, Application US/08308461
Patent No. 5706640
GENERAL INFORMATION:
APPLICANT: Jeffrey W. Voss
APPLICANT: Connie Caron
TITLE OF INVENTION: Inducers of Gamma Globin Gene Expression
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,461
FILING DATE: 16-SEP-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: oligonucleotide
US-08-308-461-3

Query Match 72.7%; Score 16; DB 1; Length 33;

Best Local Similarity 100.0%; Pred. No. 14;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TTCACTCGCCGTACT 20
|||||
Db 1 TTCACTCGCCGTACT 16

RESULT 12

PCT-US95-11823-3

Sequence 3, Application PC/US9511823

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Inducers of Gamma Globin Gene Expression

TITLE OF INVENTION: and Screening Assays Therefor

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lahive & Cockfield

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/11823

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/308,461

FILING DATE: 16-SEP-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: DeConti, Giulio A. Jr.

REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: BBI-015PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: oligonucleotide

PCT-US95-11823-3

Query Match 72.7%; Score 16; DB 5; Length 33;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TTCACTCGCCGTACT 20
|||||
Db 1 TTCACTCGCCGTACT 16

RESULT 13

US-09-106-464-1/c

Sequence 1, Application US/09106464

PATENT NO. 6011145

GENERAL INFORMATION:

APPLICANT: Steffens, John C.

APPLICANT: Shengas, Gurdev S.

APPLICANT: Kral, Jian-Ping

APPLICANT: Bannetta, Nancy

TITLE OF INVENTION: Chain length specific UDP-Glc: fatty Acid

TITLE OF INVENTION: Glucosyltransferases

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jones, Tullar & Cooper, P.C.

STREET: P.O. Box 2266 Eads Station

CITY: Arlington

STATE: Virginia

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/106,464

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/055,554

FILING DATE: 13-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: Spector, Eric S.

REGISTRATION NUMBER: 22495

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-415-1500

TELEFAX: 703-415-1508

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1627 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1413

US-09-106-464-1

Query Match

Best Local Similarity 70.0%; Score 15.4; DB 3; Length 1627;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCACCTCGCCGTACT 20
|||||
Db 1291 CTCACCTCGCCGTACT 1275

RESULT 14

US-08-778-912A-9

Sequence 9, Application US/08778912A

PATENT NO. 5876977

GENERAL INFORMATION:

APPLICANT: Wang, Jun

APPLICANT: Ngan, P N

APPLICANT: But, P P

APPLICANT: Shaw, P C

TITLE OF INVENTION: POLYMERASE CHAIN REACTION-RESTRICTION FRAGMENT

TITLE OF INVENTION: LENGTH POLYMORPHISM TEST FOR THE AUTHENTICATION OF

TITLE OF INVENTION: TRADITIONAL CHINESE MEDICINES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/778,912A

FILED DATE: 435
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chan, Albert Wai-Kit
REGISTRATION NUMBER: 36,479
REFERENCE/DOCKET NUMBER: 52188
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 278 0400
TELEFAX: 212 391 0526
INFORMATION FOR SEQ. ID NO.: 9
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-778-912A-9

Query Match 68.2%; Score 15; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ACTGCCGCTTACTAG 22
|||||
DB 1 ACTGCCGCTTACTAG 15

RESULT 15

US-08-778-912A-11
Sequence 11, Application US/0878912A
Patent No. 5876577
GENERAL INFORMATION:
APPLICANT: Wang, Jun
APPLICANT: Ngan, F N
APPLICANT: But, P P
APPLICANT: Shaw, P C
TITLE OF INVENTION: POLYMERASE CHAIN REACTION-RESTRICTION FRAGMENT
TITLE OF INVENTION: LENGTH POLYMORPHISM TEST FOR THE AUTHENTICATION OF
TITLE OF INVENTION: TRADITIONAL CHINESE MEDICINES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,912A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chan, Albert Wai-Kit
REGISTRATION NUMBER: 36,479
REFERENCE/DOCKET NUMBER: 52188
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 278 0400
TELEFAX: 212 391 0526
INFORMATION FOR SEQ. ID NO.: 11
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-778-912A-11

Query Match

68.2%; Score 15; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ACTGCCGCTTACTAG 22
|||||
DB 1 ACTGCCGCTTACTAG 15

Search completed: October 17, 2002, 08:51:48
Job time : 33.2889 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: October 17, 2002, 08:24:35 ; Search time 1156.22 Seconds
(without alignments)
256.813 Million cell updates/sec

Title: US-09-780-113D-8

Perfect score: 22

Sequence: 1 ccgcctcactgcgccttactag 22

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_hiv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	309	9	AI904541 PM-BT057-
C 2	22	100.0	311	9	AW724745 f7h10nm.f
C 3	22	100.0	312	9	AI320035
C 4	22	100.0	313	10	BE776149
C 5	22	100.0	314	9	AI904502
C 6	22	100.0	359	9	AM981798
C 7	22	100.0	359	9	AM982034
C 8	22	100.0	367	9	AM982034
C 9	22	100.0	517	10	BE356832
C 10	21	95.5	206	12	AZ925177
C 11	21	95.5	232	12	AZ917057
C 12	21	95.5	262	12	AZ932421
C 13	21	95.5	273	12	AZ932421
C 14	21	95.5	319	12	AZ916987
C 15	21	95.5	330	12	AZ124480
C 16	21	95.5	369	12	AZ925656
C 17	21	95.5	400	12	AZ924119

18	21	95.5	427	10	T39104	T39104 EST104675 S
19	21	95.5	432	12	AZ923503	AZ923503 4908..4121
20	21	95.5	437	12	AZ917126	AZ917126 4911..4464
21	21	95.5	454	12	CNS07850	AL437434 17 end of
22	21	95.5	459	12	AQ501835	AQ501835 V13H11 mt
23	21	95.5	469	12	AQ875401	AQ875401 V124H8 mt
24	21	95.5	482	12	AQ875198	AQ875198 V12366 mt
25	21	95.5	488	12	AZ923389	AZ923389 4908..4120
26	21	95.5	493	12	AQ874532	AQ874532 V110B9 mt
27	21	95.5	506	12	AZ930728	AZ930728 474..dh/57
28	21	95.5	514	12	AZ931356	AZ931356 474..dh/64
29	21	95.5	516	12	AZ928557	AZ928557 479..d1111
30	21	95.5	516	12	AZ928557	AZ928557 479..d1111
31	21	95.5	517	12	AZ928988	AZ928988 479..d1120
32	21	95.5	518	12	AZ929426	AZ929426 479..d1126
33	21	95.5	519	12	AZ928336	AZ928336 479..d1108
34	21	95.5	519	12	AZ928515	AZ928515 479..d1110
35	21	95.5	519	12	AZ929334	AZ929334 479..d1125
36	21	95.5	519	12	AZ929334	AZ929334 479..d1125
37	21	95.5	520	12	AZ928166	AZ928166 479..d1105
38	21	95.5	520	12	AZ928211	AZ928211 479..d1105
39	21	95.5	520	12	AZ929053	AZ929053 479..d1105
40	21	95.5	525	12	AQ875993	AQ875993 V131H9 mt
41	21	95.5	532	12	CNS0684P	AL394655 13 end of
42	21	95.5	542	12	AZ929712	AZ929712 479..d1188
43	21	95.5	542	12	AQ502375	AQ502375 V2764 mt
44	21	95.5	654	12	AQ492180	AQ492180 V104H11 m
45	21	95.5	664	12	CNS07ZAPC	AL436874 13 end of

ALIGNMENTS

RESULT 1
AI904541/c 309 bp mRNA linear EST 30-MAR-2000
LOCUS PM-BT057-290199-305 BT057 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 309)
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,
Nagai M.A., de Silva W. Jr., Zago M.A., Bordin S., Costa I.F.,
Goldman G.H., Carvalho A.F., Matsukuma A., Bala G.S., Simpson L.H.,
Brunstein A., de Oliveira P.S., Bucher P., Jongeneel C.V., Thare
'M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and
Simpson A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?tl=PM62-PM-BT057-305.html
33-290199&tl=1)
Seq primer: puc 18 forward.

FEATURES
source
1..309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT057"

/sex="female"
/dev stage="Adult"
/note="Organ: breast; Vector: puc18; Site:1: Sma1; Site:2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 87 a 67 c 90 g 65 t
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 88 CCGCTTCACTGCGCGTTACTAG 67

RESULT 2
LOCUS AM724745 311 bp mRNA linear EST 19-APR-2000
DEFINITION f7h10m.f1 Neurospora crassa morning cDNA library Neurospora crassa
ACCESSION AM724745
VERSION AM724745.1 GI:7619305
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 311)
AUTHORS Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
TITLE Two Neurospora crassa EST Databases
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: Universal Forward Primer
High quality sequence stop: 266.

FEATURES
Location/Qualifiers
1..311
source

/organism="Neurospora crassa"
/strain="bd, frg7 A"
/db_xref="taxon:5141"
/clone="f7h10m"
/clone_lib="Neurospora crassa morning cDNA library"
/tissue_type="tissue harvested following 22hr growth in dark"
/note="Vector: pBluescript SK-; Site:1: Xba1; Site:2: EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.
5' end of cDNA cloned into Xba1 site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 61 a 100 c 69 g 81 t
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 311;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 158 CCGCTTCACTGCGCGTTACTAG 179

RESULT 3
LOCUS AI320035 312 bp mRNA linear EST 18-DEC-1998
DEFINITION b8902nm.f1 Neurospora crassa morning cDNA library Neurospora crassa
ACCESSION AI320035
VERSION AI320035.1 GI:4036017
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 312)
AUTHORS Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
TITLE Two Neurospora crassa EST Databases
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: Universal Forward Primer
High quality sequence stop: 262.

FEATURES
Location/Qualifiers
1..312
source

/organism="Neurospora crassa"
/strain="bd, frg7 A"
/db_xref="taxon:5141"
/clone="b8902nm"
/clone_lib="Neurospora crassa morning cDNA library"
/tissue_type="tissue harvested following 22hr growth in dark"
/note="Vector: pBluescript SK-; Site:1: Xba1; Site:2: EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.
5' end of cDNA cloned into Xba1 site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 63 a 101 c 67 g 81 t
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 158 CCGCTTCACTGCGCGTTACTAG 179

RESULT 4
LOCUS BE776149/c 313 bp mRNA linear EST 20-SEP-2000
DEFINITION MY-11-E-03 Pinfeistansmy Phytophthora infestans cDNA, mRNA sequence.
ACCESSION BE776149
VERSION BE776149.1 GI:10229804
KEYWORDS EST.
SOURCE potato late blight agent.
ORGANISM Phytophthora infestans
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.

REFERENCE 1 (bases 1 to 313)
AUTHORS Kamoun,S., Hrabert,P.T., Sobral,B.W.S., Nuss,D. and Govers,F.
TITLE Initial assessment of gene diversity for the oomycete pathogen
Phytophthora infestans based on expressed sequences
JOURNAL Fungal Genet. Biol. 28 (2), 94-106 (1999)
MEDLINE 20056376
COMMENT Contact: Govers F
Laboratory of Phytopathology
Wageningen University
Bilthoven 9, P.O.Box 8025, 6700 EE, Wageningen, The Netherlands

Oct 21 16:44:31 2002

tel: 31 317 483 138
fax: 31 317 483 412
Email: Francine.Covers@medew.fylo.wau.nl.
Location/Qualifiers
1. .313
/organism="phytophthora infestans"
/strain="DDR602, Al mating type"
/db_xref="taxon:4787"
/clone_id="PinfestansW"
/dev_stage="4-week old vegetative, non-sporulating
mycelium in synthetic medium"
/lab_host="E. coli, strain DH5-alpha"
/note="Vector: pSPORT1, Site_1: SalI; Site_2: NotI; Total
kna was isolated from mycelium of P. infestans DDR602
cultured for 4 weeks in synthetic medium. EST clones were
named by their position in the microtiter plate, preceded
by the prefix MY (for mycelial) and the successive number
of the microtiter plate (e.g. MY-06-A-04)."
59 c 92 g 79 t 1 others

BASE COUNT 82 a 100.0% Score 22; DB 10; Length 313; Gaps 0;
G1N 100.0% Pred. No. 2.2; Indels 0;
Query Match 100.0% 0; Mismatches 0;
Best Local Similarity 100.0% 0; Mismatches 0;
Matches 22; Conservative 22

1 CCGCTTCACTCCGCTTACTAG 55
|||||
76 CCGCTTCACTCCGCTTACTAG 55

314 bp mRNA linear EST 30-MAR-2000
PM-BT057-100299-347 BT057 Homo sapiens cDNA, mRNA sequence.

FEATURES
source
1. .314
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="BT057"
/ov="female"

REFERENCE
AUTHORS
ORGANISM
SOURCE
KEYWORDS
VERSION
ACCESSION
DEFINITION
LOCUS
Locus 5
1904502
PM-BT057-100299-347 BT057 Homo sapiens cDNA, mRNA sequence.

TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source
1. .314
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="BT057"
/ov="female"

Smal: A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription was performed under low
stringency conditions.
88 t
91 c 63 g 88 t

BASE COUNT 72 a 100.0% Score 22; DB 9; Length 314; Gaps 0;
G1N 100.0% Pred. No. 2.2; Indels 0;
Query Match 100.0% 0; Mismatches 0;
Best Local Similarity 100.0% 0; Mismatches 0;
Matches 22; Conservative 22

1 CCGCTTCACTCCGCTTACTAG 22
|||||
222 CCGCTTCACTCCGCTTACTAG 243

359 bp mRNA linear EST 02-JUN-2000
PM-BT057-100299-347 BT057 Homo sapiens cDNA, mRNA sequence.

FEATURES
source
1. .314
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="BT057"
/ov="female"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 359)
Whetten, R.M., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
The pine gene discovery project
Unpublished (1999)
Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University, 6113 Jordan Hall, Raleigh
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
NC, 27695-8008
Tel: 919-515-7801
Fax: 919-515-7801
Email: rosswhetten@unity.ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing primer.

1. .359
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone_id="Pinus taeda"
/dev_stage="Immature"
/lab_host="E. coli BM25.8"
/note="Organ: Pollen cone; Vector: Lambda triplex; Site_1:
Site_2: SfiI (B); Immature pollen cones were
collected in the early spring, frozen and used for RNA
isolation. The SMART-PCR method (Clontech) was used to
prepare a library from 1 ug total RNA, using the Lambda
Triplex vector. Plasmid subclones in E. coli strain BM25.8 and
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
74 c 101 g 83 t

BASE COUNT 101 a 100.0% Score 22; DB 9; Length 359; Gaps 0;
G1N 100.0% Pred. No. 2.3; Indels 0;
Query Match 100.0% 0; Mismatches 0;
Best Local Similarity 100.0% 0; Mismatches 0;
Matches 22; Conservative 22

1 CCGCTTCACTCCGCTTACTAG 22
|||||
103 CCGCTTCACTCCGCTTACTAG 82

RESULT 7
AM982034/c

Mon Oct 21 16:44:31 2002

US-09-780-113d-8.1st

LOCUS AM982034
DEFINITION PC23A10 Pine Triplex pollen cone library pinus taeda cDNA clone
ACCESSION PC23A10 Pine Triplex pollen cone library pinus taeda cDNA clone
VERSION AM982034 mRNA sequence.
KEYWORDS AM982034.1 GI:8173602
SOURCE EST
ORGANISM 10b1011y pine.
pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Coniferophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
JOURNAL The Pine Gene Discovery Project
COMMENT Unpublished (1999)
Contact: Ross Whetten
Forest Biotechnology
North Carolina State University
Dept. of Forestry, NC State University
Tel: 27695-8008
Fax: 919-515-7800
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.

FEATURES
Source
Location/Qualifiers
1..359
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC23A10"
/dev_stage="Immature"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1;
Isolation. The SMART-PCR method (Clontech) was used for mRNA
prepare a library from 1 ug total RNA, using the lambda
Triplex vector. Plasmid subclones in pTriplex were
recovered by Cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end." 83 t
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 359;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
103 CCGCTTCACTGCGCGTTACTAG 22
|||||
103 CCGCTTCACTGCGCGTTACTAG 22

BASE COUNT
ORIGIN 101 a
73 c 102 g 83 t
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 359;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
103 CCGCTTCACTGCGCGTTACTAG 22
|||||
103 CCGCTTCACTGCGCGTTACTAG 22

FEATURES
Source
Location/Qualifiers
1..367
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC03D12"
/dev_stage="Immature"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1;
Isolation. The SMART-PCR method (Clontech) was used for mRNA
prepare a library from 1 ug total RNA, using the lambda
Triplex vector. Plasmid subclones in pTriplex were
recovered by Cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end." 84 t
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 367;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
103 CCGCTTCACTGCGCGTTACTAG 22
|||||
103 CCGCTTCACTGCGCGTTACTAG 22

BASE COUNT
ORIGIN 99 a
70 c 105 g 84 t
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 367;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
103 CCGCTTCACTGCGCGTTACTAG 22
|||||
103 CCGCTTCACTGCGCGTTACTAG 22

DEFINITION DC1_145.D09.bl.A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
LOCUS BE356832/c
ACCESION BE356832
VERSION BE356832
KEYWORDS EST
SOURCE EST: 9298389
ORGANISM Sorghum.
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
REFERENCE 1 (bases 1 to 517)
AUTHORS Cordonnier-Pratt, M.W., Gingle, A., Marsala, C., Sudman, M. and Pratt
TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector and regions
below phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 418
POLY-A-NO.
Location/Qualifiers
1..517
/organism="Sorghum"

Query Match 100.0%; Score 22; DB 10; Length 517;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACCTCGCGCTTACTAG 22
|||||
Db 75 CCGCTTCACCTCGCGCTTACTAG 54

RESULT 10
AZ925177 206 bp DNA linear GSS 01-APR-2001
LOCUS 4910.ez32908.s1 Saccharomyces paradoxus N17 Saccharomyces paradoxus
DEFINITION genomic clone 4910.ez32908.s1, DNA sequence.
ACCESSION AZ925177
VERSION AZ925177.1 GI:13496076
KEYWORDS GSS.
SOURCE Saccharomyces paradoxus.
ORGANISM Saccharomyces paradoxus.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 206)
AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
,W.R., Waterston,R.H. and Johnston,M.
TITLE Surveying Saccharomyces genomes to identify functional elements by
comparative DNA sequence analysis
JOURNAL Unpublished (2001)
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mj@genetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..206
/organism="Saccharomyces paradoxus"
/strain="N17"
/db_xref="taxon:27291"
/clone="4910.ez32908.s1"
/clone_lib="Saccharomyces paradoxus N17"
/note="Random genomic sequence"

BASE COUNT 51 a 64 c 39 g 52 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACCTCGCGCTTACTA 21
|||||
Db 164 CCGCTTCACCTCGCGCTTACTA 184

RESULT 11
AZ917057 232 bp DNA linear GSS 01-APR-2001
LOCUS 4911.f663n01.s1 Saccharomyces bayanus MCYC 623-6C Saccharomyces
DEFINITION bayanus genomic clone 4911.f663n01.s1, DNA sequence.
ACCESSION AZ917057
VERSION AZ917057.1 GI:13493581
KEYWORDS GSS.
SOURCE Saccharomyces bayanus.
ORGANISM Saccharomyces bayanus.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 232)
AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
,W.R., Waterston,R.H. and Johnston,M.
TITLE Surveying Saccharomyces genomes to identify functional elements by
comparative DNA sequence analysis

JOURNAL Unpublished (2001)
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mj@genetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..232
/organism="Saccharomyces bayanus"
/strain="MCYC 623-6C"
/db_xref="taxon:4931"
/clone="4911.f663n01.s1"
/clone_lib="Saccharomyces bayanus MCYC 623-6C"
/note="Random genomic sequence"

BASE COUNT 68 a 44 c 62 g 58 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 232;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACCTCGCGCTTACTA 21
|||||
Db 112 CCGCTTCACCTCGCGCTTACTA 92

RESULT 12
AZ932421 262 bp DNA linear GSS 01-APR-2001
LOCUS 474.dhz9g11.s1 Saccharomyces unisporus NRRL Y-1556 Saccharomyces
DEFINITION unisporus genomic clone 474.dhz9g11.s1, DNA sequence.
ACCESSION AZ932421
VERSION AZ932421.1 GI:13503334
KEYWORDS GSS.
SOURCE Saccharomyces unisporus.
ORGANISM Saccharomyces unisporus.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 262)
AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
,W.R., Waterston,R.H. and Johnston,M.
TITLE Surveying Saccharomyces genomes to identify functional elements by
comparative DNA sequence analysis
JOURNAL Unpublished (2001)
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mj@genetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..262
/organism="Saccharomyces unisporus"
/strain="NRRL Y-1556 (CBS 398)"
/db_xref="taxon:27294"
/clone="474.dhz9g11.s1"
/clone_lib="Saccharomyces unisporus NRRL Y-1556"
/note="Random genomic sequence"

BASE COUNT 69 a 84 c 43 g 66 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 262;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACCTCGCGCTTACTA 21
|||||
Db 211 CCGCTTCACCTCGCGCTTACTA 231

RESULT 13
LOCUS A2932397 273 bp DNA linear GSS 01-APR-2001
DEFINITION 474.dhz98e09.s1 Saccharomyces unisporus NRRL Y-1556 Saccharomyces unisporus genomic clone 474.dhz98e09.s1, DNA sequence.
ACCESSION A2932397
KEYWORDS A2932397.1 GI:13503310
SOURCE GSS.
ORGANISM Saccharomyces unisporus.
Saccharomyces unisporus.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 273)
Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.
Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis
Unpublished (2001)
JOURNAL
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mjgenetics.wustl.edu
Class: random plasmid subclone.
FEATURES
Location/Qualifiers
1..273
/organism="Saccharomyces unisporus"
/strain="NRRL Y-1556 (CBS 398)"
/db_xref="taxon:27294"
/clone="474.dhz98e09.s1"
/clone_lib="Saccharomyces unisporus NRRL Y-1556"
/note="Random genomic sequence"

BASE COUNT 64 a 78 c 60 g 71 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTA 21
|||||
Db 13 CCGCTTCACTCGCGCTTACTA 33

RESULT 14
LOCUS A2916987 319 bp DNA linear GSS 01-APR-2001
DEFINITION 4911.f663j12.s1 Saccharomyces bayanus MCYC 623-6C Saccharomyces bayanus genomic clone 4911.f663j12.s1, DNA sequence.
ACCESSION A2916987
KEYWORDS A2916987.1 GI:13493511
SOURCE GSS.
ORGANISM Saccharomyces bayanus.
Saccharomyces bayanus.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 319)
Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.
Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis
Unpublished (2001)
JOURNAL
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mjgenetics.wustl.edu

Class: random plasmid subclone.
FEATURES
Location/Qualifiers
1..319
/organism="Saccharomyces bayanus"
/strain="MCYC 623-6C"
/db_xref="taxon:4931"
/clone="4911.f663j12.s1"
/clone_lib="Saccharomyces bayanus MCYC 623-6C"
/note="Random genomic sequence"

BASE COUNT 91 a 57 c 92 g 79 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 319;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTA 21
|||||
Db 145 CCGCTTCACTCGCGCTTACTA 125

RESULT 15
LOCUS A2124480 330 bp DNA linear GSS 01-JUN-2000
DEFINITION K025r S. Kluyveri genomic library Saccharomyces kluyveri genomic, DNA sequence.
ACCESSION A2124480
VERSION A2124480.1 GI:8889034
KEYWORDS GSS.
SOURCE Saccharomyces kluyveri.
ORGANISM Saccharomyces kluyveri.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 330)
Langkjaer,R.B., Nielsen,M.L., Daugaard,P.R., Liu,W. and Piskur,J.
Yeast chromosomes have been significantly reshaped during their evolutionary history
Unpublished (2000)
JOURNAL
COMMENT Contact: Langkjaer RB and Piskur J
Department of Microbiology
Technical University of Denmark
Building 301, DK-2800 Lyngby, Denmark
Tel: 0045 45 45 25 18
Fax: 0045 45 93 28 09
Email: imbl@pop.dtu.dk and imj@pop.dtu.dk
Class: plasmid ends.
FEATURES
Location/Qualifiers
1..330
/organism="Saccharomyces kluyveri"
/strain="NRRL Y-12651r"
/db_xref="taxon:4934"
/clone_lib="S. Kluyveri genomic library"
/note="Vector: pBR322; Partial Sau3A fragments in BamHI"

BASE COUNT 83 a 85 c 69 g 93 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTA 21
|||||
Db 58 CCGCTTCACTCGCGCTTACTA 78

Search completed: October 17, 2002, 10:03:52
Job time : 1161.22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 07:07:41 : Search time 32.7111 Seconds
(without alignments)
172.711 Million cell updates/sec

Title: US-09-780-113D-15

Perfect score: 23

Sequence: 1 accacgactgagcagcactt 23

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/lna/PCUTUS.COMB.seq:*

6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.6	72.2	1181	2 US-08-632-598-2	Sequence 2, Appl1
2	16.6	72.2	1181	4 US-09-231-240-2	Sequence 2, Appl1
3	16.6	72.2	10079	2 US-08-476-866-20	Sequence 20, Appl1
4	15.6	67.8	686	4 US-08-943-731-116	Sequence 116, App
5	15.6	67.8	24183	4 US-08-943-731-3	Sequence 3, Appl1
6	15.2	66.1	821	4 US-08-735-836C-136	Sequence 136, App
7	15.2	66.1	2246	4 US-09-117-860-17	Sequence 17, Appl1
8	15.2	66.1	3435	1 US-08-366-577-1	Sequence 1, Appl1
9	15.2	66.1	3435	1 PCT-US96-00005-1	Sequence 1, Appl1
10	15	65.2	11219	3 US-07-642-734C-1	Sequence 1, Appl1
11	15	65.2	11219	3 US-08-439-009A-1	Sequence 1, Appl1
12	15	65.2	51259	4 US-08-781-891-209	Sequence 209, App
13	14.8	64.3	1854	4 US-08-943-731-650	Sequence 650, App
14	14.8	64.3	1854	4 US-09-230-944-9	Sequence 9, Appl1
15	14.8	64.3	1874	1 US-08-779-700-3	Sequence 3, Appl1
16	14.8	64.3	1919	1 US-08-779-700-5	Sequence 5, Appl1
17	14.8	64.3	1919	1 US-08-779-700-7	Sequence 7, Appl1
18	14.8	64.3	1919	1 US-08-779-700-11	Sequence 11, Appl1
19	14.6	63.5	1778	1 PCT-US91-02958-9	Sequence 9, Appl1
20	14.6	63.5	3348	1 US-08-222-616-35	Sequence 35, Appl1
21	14.6	63.5	3348	5 PCT-US95-04228-35	Sequence 35, Appl1
22	14.6	63.5	5761	1 US-07-749-001-2	Sequence 2, Appl1
23	14.6	63.5	5761	1 US-08-154-198-2	Sequence 2, Appl1
24	14.6	63.5	5761	1 US-08-463-335-2	Sequence 2, Appl1
25	14.6	63.5	5761	2 US-08-464-023A-2	Sequence 2, Appl1
26	14.6	63.5	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
27	14.2	61.7	880	1 US-08-616-368A-7	Sequence 7, Appl1

ALIGNMENTS

C 28	14.2	61.7	880	3	US-09-054-298-7	Sequence 7, Appl1
C 29	14.2	61.7	880	4	US-08-818-653-7	Sequence 7, Appl1
C 30	14.2	61.7	1130	4	US-08-936-165A-248	Sequence 248, App
C 31	14.2	61.7	1652	1	US-08-036-555B-148	Sequence 148, App
C 32	14.2	61.7	1652	1	US-08-469-569-148	Sequence 148, App
C 33	14.2	61.7	1652	1	US-08-249-322A-148	Sequence 148, App
C 34	14.2	61.7	1652	1	US-08-469-526A-148	Sequence 148, App
C 35	14.2	61.7	1652	2	US-08-734-591A-148	Sequence 148, App
C 36	14.2	61.7	1652	2	US-08-469-660-148	Sequence 148, App
C 37	14.2	61.7	1652	3	US-08-341-018-53	Sequence 53, Appl1
C 38	14.2	61.7	1652	3	US-08-470-335-148	Sequence 148, App
C 39	14.2	61.7	1652	4	US-08-735-021-148	Sequence 148, App
C 40	14.2	61.7	1652	4	US-08-734-664A-148	Sequence 148, App
C 41	14.2	61.7	1652	4	US-08-470-339-148	Sequence 148, App
C 42	14.2	61.7	1652	5	PCT-US94-05083C-144	Sequence 144, App
C 43	14.2	61.7	1652	5	PCT-US95-06846A-148	Sequence 148, App
C 44	14.2	61.7	16075	3	US-09-096-942-1	Sequence 1, Appl1
C 45	14.2	61.7	16075	3	US-09-096-867-1	Sequence 1, Appl1

RESULT 1

US-08-632-598-2

Sequence 2, Application US/08632598

Patent No. 586164

GENERAL INFORMATION:

APPLICANT: BIRD, COLIN R

TITLE OF INVENTION: RIPENING-RELATED GENES FROM BANANA

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS: CUSHMAN DARBY AND CUSHMAN

STREET: 1100 NEW YORK AVENUE N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3918

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/632,598

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOKUTS, PAUL N.

REGISTRATION NUMBER: 16,773

REFERENCE/DOCKET NUMBER: 223355/SEE50112/US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 861-3000

TELEFAX: 822-0944

TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1181 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

ORGANISM: MOSA

IMMEDIATE SOURCE:

CLONE: EPE GENE

US-08-632-598-2

Query Match 72.2% Score 16.6; DB 2; Length 1181;
Best Local Similarity 82.6% Pred. No. 13;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCACGACTGACGACGACCTT 23
||| ||||| ||| |||||
DB 289 ACCTGACTGGAGGACGACCTT 311

RESULT 2

US-09-231-240-2
Sequence 2, Application US/09231240
Patent No. 6262346
GENERAL INFORMATION:
APPLICANT: BIRD, COLIN R
APPLICANT: FLETCHER, JONATHAN D
TITLE OF INVENTION: RIPENING-RELATED GENES FROM BANANA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DAREY AND CUSHMAN
STREET: 1100 NEW YORK AVENUE N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,240
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,598
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 223355/SEE50112/US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 861-3000
TELEFAX: 822-0944
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: MUSA
IMMEDIATE SOURCE:
CLONE: EFE GENE
US-09-231-240-2

Query Match 72.2%; Score 16.6; DB 4; Length 1181;
Best Local Similarity 82.6%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCACGACTGACGACGACCTT 23
||| ||||| ||| |||||
DB 289 ACCTGACTGGAGGACGACCTT 311

RESULT 3

US-08-476-866-20/c
Sequence 20, Application US/08476866
Patent No. 5994359

GENERAL INFORMATION:
APPLICANT: GRAP, JAMES D.
APPLICANT: FRIDOVICH, IRWIN
APPLICANT: OURY, TIM
APPLICANT: DAY, BRIAN J.
APPLICANT: FOLZ, RODNEY J.

APPLICANT: FREEMAN, BRUCE A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE AND MIMETICS THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,866
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/322,766
FILING DATE: 13-OCT-1994
APPLICATION NUMBER: US 08/136,207
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-74
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 10079 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 5086..5803
US-08-476-866-20

Query Match 72.2%; Score 16.6; DB 2; Length 10079;
Best Local Similarity 82.6%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCACGACTGACGACGACCTT 23
||| ||||| ||| |||||
DB 1238 ACCTGACTGAGGACGACCTT 1216

RESULT 4

US-08-943-731-116/c
Sequence 116, Application US/08943731
Patent No. 6265157

GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORKKO, JARMO
APPLICANT: ALA-KOKKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:

ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 686 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-116

Query Match 67.8%; Score 15.6; DB 4; Length 686;
Best Local Similarity 81.8%; Pred. No. 44;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCAGCAGTGAGCAGCACCTT 23
Db 408 CGAGACTGAGCAGCACCTCT 387

RESULT 5
US-08-943-731-3/C
Sequence 3, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA

ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24183 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-3

Query Match 67.8%; Score 15.6; DB 4; Length 24183;
Best Local Similarity 81.8%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCAGCAGTGAGCAGCACCTT 23
Db 4798 CGAGACTGAGCAGCACCTCT 4777

RESULT 6
US-08-235-836C-136
Sequence 136, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BN193-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-235-836C-136

Query Match 66.1%; Score 15.2; DB 4; Length 821;
Best Local Similarity 85.0%; Pred. No. 72;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

UY 1 ACCAGCTGAGCAGCAGC 20
Db 740 ACTAGCTGAGCAGCAGC 759

RESULT 7
US-09-117-860-17
Sequence 17, Application US/09117860A
Patent No. 6338955
GENERAL INFORMATION:
APPLICANT: OGURI, Suguru
APPLICANT: MINOWA, Mari
APPLICANT: YOSHIDA, Aruto
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: TAKEUCHI, Makoto
TITLE OF INVENTION: NOVEL 1-4-N-ACETYLGLUCOSAMINYLTRANSFERASE AND GENE
TITLE OF INVENTION: ENCODING THE SAME
FILE REFERENCE: 081356/0119
CURRENT APPLICATION NUMBER: US/09/117,860A
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: WO PCT/JP97/04546
EARLIER FILING DATE: 1997-12-10
EARLIER APPLICATION NUMBER: JP 161462/1997
EARLIER FILING DATE: 1997-06-18
EARLIER APPLICATION NUMBER: JP 332411/1996
EARLIER FILING DATE: 1996-12-12
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 17
LENGTH: 2246
TYPE: DNA
ORGANISM: Bovine
FEATURE:
NAME/KEY: CDS
LOCATION: (288)..(1892)
US-09-117-860-17

Query Match 66.1%; Score 15.2; DB 4; Length 2246;
Best Local Similarity 85.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

UY 4 ACAGCTGAGCAGCAGCTT 23
Db 1950 ACAGCTGAGCAGCAGCTT 1969

RESULT 8
US-08-366-577-1/c
Sequence 1, Application US/08366577
Patent No. 5728523
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
TITLE OF INVENTION: POLYMERASE DELTA MUTATIONS IN COLORECTAL
TITLE OF INVENTION: TUMORS WITH REPLICATION ERRORS
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/366,577
FILING DATE: 12-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107,48554
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3435 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 43..3364
US-08-366-577-1

Query Match 66.1%; Score 15.2; DB 1; Length 3435;
Best Local Similarity 85.0%; Pred. No. 88;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

UY 2 CCAGACTGAGCAGCAGCT 21
Db 1691 CCAGACTGAGCAGCAGCT 1672

RESULT 9
PCT-US96-00005-1/c
Sequence 1, Application PC/T095600005
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
TITLE OF INVENTION: POLYMERASE DELTA MUTATIONS IN COLORECTAL
TITLE OF INVENTION: TUMORS WITH REPLICATION ERRORS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00005
FILING DATE: 2-JAN-96

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.53505
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMD UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3435 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 43..3364
PCT-US96-00005-1

Query Match 66.1%; Score 15.2; DB 5; Length 3435;
Best Local Similarity 85.0%; Pred. No. 88;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGACTGACGACGACCT 21
Db 1691 CCAGACTGACGACGACT 1672

RESULT 10
US-07-642-734C-1/C
Sequence 1, Application US/07642734C
Patent No. 5824513
GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/AF6D-2 One Abbott
STREET: Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andrew M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11219 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 2338
FEATURE:
NAME/KEY: misc.feature
LOCATION: 744..6659
OTHER INFORMATION: /function= "APPROXIMATE SPAN OF
MODULE 1"
OTHER INFORMATION: /label= FUNCTION
FEATURE:
NAME/KEY: CDS
LOCATION: 744..11219
OTHER INFORMATION: /function= "gene- "erya""
OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for
OTHER INFORMATION: 6-deoxyerythronolide B""
FEATURE:
NAME/KEY: misc.feature
LOCATION: 744..1868
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain 1 of module 1"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1998..2198
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 1 of module 1"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 2250..3626
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3831..4811
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase 2 domain of module 1"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 5574..6125
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 1"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 6369..6626
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 2 of module 1"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 6678..11219
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 2"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 6678..8066
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 2"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 8262..9305
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 2"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 9906..10454
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 2"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 10707..10964
```

```

; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 2"
US-07-642-734C-1

Query Match          65.2%, Score 15; DB 1; Length 11219;
Best Local Similarity 78.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCTT 23
Db 4333 ACCACGACGAGCGGACCGTT 4311

RESULT 11
US-08-439-009A-1/C
; Sequence 1, Application US/08439009A
; Patent No. 6004787
; GENERAL INFORMATION:
; APPLICANT: Donadio, S
; APPLICANT: Katz, L
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Method of Directing Biosynthesis of
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven F. Weinstein
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,009A
; FILING DATE: 11-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Dianne
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 4952-US.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11219 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharopolyspora erythraea
; STRAIN: NRRL 2338
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 744..6659
; OTHER INFORMATION: /function= "APPROXIMATE SPAN OF
; OTHER INFORMATION: MODULE 1"
; OTHER INFORMATION: /label= FUNCTION
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 744..11219
; OTHER INFORMATION: /function= "gene= "eryA"
; OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for
; OTHER INFORMATION: 6-deoxyerythronolide B"
; FEATURE:
; NAME/KEY: misc_feature
```

```

; LOCATION: 744..1868
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain 1 of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1998..2198
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain 1 of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2250..3626
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3831..4811
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase 2 domain of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5574..6125
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6369..6626
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain 2 of module 1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6678..11219
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6678..8066
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoacyl ACP synthase of module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8262..9305
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyltransferase domain of module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9906..10454
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: beta-ketoreductase domain of module 2"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10707..10964
; OTHER INFORMATION: /function= "approximate span of
; OTHER INFORMATION: acyl carrier domain of module 2"
US-08-439-009A-1

Query Match          65.2%, Score 15; DB 3; Length 11219;
Best Local Similarity 78.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCTT 23
Db 4333 ACCACGACGAGCGGACCGTT 4311

RESULT 12
US-08-781-891-209
; Sequence 209, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
```

APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620Leimbarg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 65.2%; Score 15; DB 3; Length 51259;
Best Local Similarity 78.3%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCAGACTAGCAGCAGCTTT 23
DB 49944 ACCATGCTGAGCAGCTCTCT 49966

RESULT 13
US-08-943-731-650
Sequence 650, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE, LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 650:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-650

Query Match 64.3%; Score 14.8; DB 4; Length 20;
Best Local Similarity 88.9%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 GACTGACGACGACCTTT 23
DB 2 GATGACGACGACCTCT 19

RESULT 14
US-09-230-944-9
Sequence 9, Application US/09230944
Patent No. 6277380
GENERAL INFORMATION:
APPLICANT: UEDA, Shigeharu
APPLICANT: WATANABE, Michiko
APPLICANT: KAWANISHI, Hitomi
TITLE OF INVENTION: Measles Virus Mutant Antigen and Gene Coding for the
FILE REFERENCE: 0216-0407P
CURRENT APPLICATION NUMBER: US/09/230,944
CURRENT FILING DATE: 1999-02-04
EARLIER APPLICATION NUMBER: PCT/JP98/02481
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 9-184285 JAPAN
EARLIER FILING DATE: 1997-06-04
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 1854
TYPE: DNA
ORGANISM: Measles virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1851)
FEATURE:
OTHER INFORMATION: Attenuated measles virus NA strain
FEATURE:
OTHER INFORMATION: any n or Xaa = Unknown
US-09-230-944-9

Query Match 64.3%; Score 14.8; DB 4; Length 1854;
Best Local Similarity 88.9%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CACGACTGAGCAGCACC 20
 Db 740 CACACTGAGCAGCACC 757

RESULT 15

US-08-279-700-3
 ; Sequence 3, Application US/08279700
 ; Patent No. 5578448
 ; GENERAL INFORMATION:
 ; APPLICANT: ROTA, Jennifer S.
 ; APPLICANT: BELINI, William J.
 ; TITLE OF INVENTION: WILD-TYPE MEASLES VIRUS GLYCOPROTEINS;
 ; TITLE OF INVENTION: VACCINE AND DETECTION METHOD THEREFOR
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentia Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/279,700
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/866,033A
 ; FILING DATE: 19920408
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 40399/140 NIHD
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ. ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1874 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 21..1874
 ; US-08-279-700-3

Query Match 64.3%; Score 14.8; DB 1; Length 1874;
 Best Local Similarity 88.9%; Pred. No. 1.3e+02;
 Matches 16; Conservative 0; Mismatches 2; Indels 0;

QY 3 CACGACTGAGCAGCACC 20
 Db 760 CACACTGAGCAGCACC 777

Search completed: October 17, 2002, 08:52:21
 Job time : 65.7111 secs

AUTHORS Brown,S., Rouse,G., Hutchings,P. and Colgan,D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6
 College Street, Sydney, NSW 2000, Australia

FEATURES
 source Location/Qualifiers

1.307
 /organism="Mesochorepterus sp. AMW22402"
 /specimen_voucher="AMW22402"
 /db_xref="taxon:104707"
 <1.>307
 /note="D1 expansion region"
 /product="28S ribosomal RNA"
 BASE COUNT 73 a 78 c 90 g 66 t
 ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 307;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTAG 22
 |||||
 Db 60 CCGCTTCACTCGCGCTTACTAG 39

Search completed: October 17, 2002, 09:24:05
 Job time : 939.733 secs

Db 58 CCGCTTACTCGCGCTTACTAG 37
 |||
 RESULT 12
 AF185171/c 305 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Polypophthalmus pictus 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185171
 VERSION AF185171.1 GI:6013352
 KEYWORDS
 SOURCE Polypophthalmus pictus.
 ORGANISM Polypophthalmus pictus.
 Eukaryota; Metazoa; Annelida; Polychaeta; Opheliida; Opheliidae;
 Polypophthalmus.
 REFERENCE 1 (bases 1 to 305)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 305)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6 College Street, Sydney, NSW 2000, Australia
 FEATURES
 SOURCE
 1. 305
 /organism="Polypophthalmus pictus"
 /specimen_voucher="AMW10095"
 /db_xref="taxon:104727"
 <1..>305
 /note="D1 expansion region"
 /product="28S ribosomal RNA"
 BASE COUNT 61 a 84 c 97 g 59 t 4 others
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 305;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTACTCGCGCTTACTAG 22
 |||
 Db 60 CCGCTTACTCGCGCTTACTAG 39
 RESULT 13
 AF185150/c 307 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Amphiglena terebro 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185150
 VERSION AF185150.1 GI:6013331
 KEYWORDS
 SOURCE Amphiglena terebro.
 ORGANISM Amphiglena terebro.
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canallipalpata;
 Sabellida; Sabellidae; Amphiglena.
 REFERENCE 1 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6 College Street, Sydney, NSW 2000, Australia
 FEATURES
 SOURCE
 1. 307
 /organism="Amphiglena terebro"
 /specimen_voucher="AMW21360"
 /db_xref="taxon:104683"
 <1..>307
 /note="D1 expansion region"

BASE COUNT 69 a 79 c 92 g 65 t 2 others
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 307;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTACTCGCGCTTACTAG 22
 |||
 Db 60 CCGCTTACTCGCGCTTACTAG 39
 RESULT 14
 AF185154/c 307 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Glycera tridactyla 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185154
 VERSION AF185154.2 GI:6073854
 KEYWORDS
 SOURCE Glycera tridactyla.
 ORGANISM Glycera tridactyla.
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculiata;
 Phyllodoctida; Glyceridae; Glycera.
 REFERENCE 1 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6 College Street, Sydney, NSW 2000, Australia
 COMMENT On Oct 20, 1999 this sequence version replaced gi:6013335.
 FEATURES
 SOURCE
 1. 307
 /organism="Glycera tridactyla"
 /specimen_voucher="AMW196835"
 /db_xref="taxon:1047710"
 <1..>307
 /note="D1 expansion region"
 /product="28S ribosomal RNA"
 BASE COUNT 62 a 81 c 97 g 67 t
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 307;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTACTCGCGCTTACTAG 22
 |||
 Db 60 CCGCTTACTCGCGCTTACTAG 39
 RESULT 15
 AF185163/c 307 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Mesochaetopterus sp. AMW22402 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185163
 VERSION AF185163.1 GI:6013344
 KEYWORDS
 SOURCE Mesochaetopterus sp. AMW22402.
 ORGANISM Mesochaetopterus sp. AMW22402.
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canallipalpata;
 Spionida; Chaetopteridae; Mesochaetopteridae.
 REFERENCE 1 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 307)

FEATURES
source
Location/Qualifiers
1.283
/organism="uncultured basidiomycete Cons3"
/isolate="Cons3"
/db_xref="taxon:150676"
/country="France: Antifer Harbour"
<1..>283
/product="24S large subunit ribosomal RNA"
78 a 53 c 89 g 63 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 22; DB 8; Length 283;
Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 46 CCGCTTCACTGCGCGTTACTAG 25

RESULT 9
AF339505/c 295 bp DNA linear INV 04-JUN-2001
LOCUS
DEFINITION
AF339505 Aplysia depilans 28S ribosomal RNA gene, partial sequence.
ACCESSION
AF339505.1 GI:13445764
VERSION
KEYWORDS
SOURCE
ORGANISM
Aplysia depilans.
Aplysia depilans
Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia;
Anaspidae; Aplysiidae; Aplysia.
1 (bases 1 to 295)
Tillier,S., Masselot,M., Guerdoux,J. and Tillier,A.
Monophyly of major Gastropod taxa tested from partial 28S rRNA
sequences, with emphasis on Euthyrea and hot vent limpets
Peltospiridae
Nautilus 108 (supplement 2), 122-140 (1994)
2 (bases 1 to 295)
Dayrat,B., Tillier,A., Lecointre,G. and Tillier,S.
New clades of euthyreaan gastropods (Mollusca) from 28S rRNA
sequences
Molecular phylogenetics and evolution. 19 (2), 225-235 (2001)
21240533
PUBMED
11341805
JOURNAL
MEDLINE
3 (bases 1 to 295)
Tillier,S., Masselot,M., Guerdoux,J. and Tillier,A.
Direct Submission
Submitted (23-JAN-2001) Laboratoire de Biologie des Invertebrés
Marins et Malacologie, Institut de Systematique FR 1541, Muséum
National d'Histoire Naturelle, 43 rue Cuvier, Paris 75005, France
JOURNAL
TITLES
FEATURES
source
Location/Qualifiers
1..295
/organism="Aplysia depilans"
/db_xref="taxon:76186"
<1..>295
/note="D1 domain"
/product="28S ribosomal RNA"
72 a 76 c 91 g 55 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 22; DB 3; Length 295;
Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 100 CCGCTTCACTGCGCGTTACTAG 79

RESULT 10
AF265622/c 303 bp DNA linear INV 09-JUL-2000
LOCUS

DEFINITION
AF265622 Madracis pharensis 28S ribosomal RNA gene, partial sequence.
AF265622.1 GI:8980711
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Madracis pharensis.
Madracis pharensis
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Astrocoeniina; Pocilloporidae; Madracis.
1 (bases 1 to 303)
Romano,S.L. and Cairns,S.D.
Molecular phylogenetic hypotheses for the evolution of
scleractinian corals
unpublished
2 (bases 1 to 303)
Romano,S.L. and Cairns,S.D.
Direct Submission
Submitted (09-MAY-2000) Laboratory of Molecular Systematics,
National Museum of Natural History, Smithsonian Institution,
Washington, DC 20560-0163, USA
JOURNAL
TITLES
REFERENCE
AUTHORS
JOURNAL
TITLES
FEATURES
source
Location/Qualifiers
1..303
/organism="Madracis pharensis"
/db_xref="taxon:123772"
<1..>303
/product="28S ribosomal RNA"
83 a 64 c 92 g 64 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 22; DB 3; Length 303;
Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 58 CCGCTTCACTGCGCGTTACTAG 37

RESULT 11
AF265642/c 304 bp DNA linear INV 09-JUL-2000
LOCUS
DEFINITION
AF265642 Caryophyllia inornata 28S ribosomal RNA gene, partial sequence.
ACCESSION
AF265642.1 GI:8980731
VERSION
KEYWORDS
SOURCE
ORGANISM
Caryophyllia inornata.
Caryophyllia inornata
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Caryophyllina; Caryophylliidae; Caryophyllia.
1 (bases 1 to 304)
Romano,S.L. and Cairns,S.D.
Molecular phylogenetic hypotheses for the evolution of
scleractinian corals
unpublished
2 (bases 1 to 304)
Romano,S.L. and Cairns,S.D.
Direct Submission
Submitted (09-MAY-2000) Laboratory of Molecular Systematics,
National Museum of Natural History, Smithsonian Institution,
Washington, DC 20560-0163, USA
JOURNAL
TITLES
REFERENCE
AUTHORS
JOURNAL
TITLES
FEATURES
source
Location/Qualifiers
1..304
/organism="Caryophyllia inornata"
/db_xref="taxon:130053"
<1..>304
/product="28S ribosomal RNA"
80 a 68 c 92 g 64 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 22; DB 3; Length 304;
Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22

VERSION U65520.1 GI:1553147
 KEYWORDS Seriatopora hystrix.
 SOURCE Seriatopora hystrix
 ORGANISM Eukaryota; Metazoa; Anthozoa; Zoantharia; Scleractinia;
 Asterozoa; Pocilloporidae; Seriatopora.
 REFERENCE 1 (bases 1 to 222)
 AUTHORS Chen, C.A., Odorico, D.M., ten Lohuis, M., Veron, J.E. and Miller, D.J.
 TITLE Systematic relationships within the Anthozoa (Cnidaria: Anthozoa)
 JOURNAL Mol. Phylogenet. Evol. 4 (2), 175-183 (1995)
 MEDLINE 95392827
 REFERENCE 2 (bases 1 to 222)
 AUTHORS Veron, J.E.N., Odorico, D.M., Chen, C.A. and Miller, D.J.
 TITLE Reassessing evolutionary relationships of scleractinian corals
 JOURNAL Coral Reefs 15, 1-9 (1996)
 REFERENCE 3 (bases 1 to 222)
 AUTHORS Odorico, D.M., Chen, C.A. and Miller, D.J.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUL-1996) Biochemistry and Molecular Biology, James Cook University, Townsville, Queensland 4811, Australia
 FEATURES
 SOURCE 1..222
 /organism="Seriatopora hystrix"
 /db_xref="taxon:51070"
 RNA 1..>222
 /product="28S ribosomal RNA"
 BASE COUNT 59 a 53 c 68 g 42 t
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 222;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTCACTGCGCGTTACTAG 22
 ||||||||||||||||||
 Db 58 CCGCTTCACTGCGCGTTACTAG 37
 RESULT 6
 TS285259/c 259 bp mRNA linear INV 20-AUG-1996
 LOCUS TS285259
 DEFINITION Tubifex sp. gene for 28S ribosomal RNA (partial: 259 bp).
 ACCESSION X90690
 VERSION X90690.1 GI:1067238
 KEYWORDS 28S ribosomal RNA; 28S rRNA gene.
 SOURCE Tubifex sp.
 ORGANISM Tubifex sp.
 Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 Tubificina; Tubificidae; Tubifex.
 REFERENCE 1 (bases 1 to 259)
 AUTHORS Friedrich, M.
 TITLE Direct Submission
 JOURNAL Submitted (09-AUG-1995) M. Friedrich, Zoological Institute, Univ. Muenchen, Luisenstr. 14, D-80333 Muenchen, FRG
 REFERENCE 2 (bases 1 to 259)
 AUTHORS Friedrich, M. and Tautz, D.
 TITLE Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods
 JOURNAL Nature 376 (6536), 165-167 (1995)
 MEDLINE 95327188
 FEATURES
 SOURCE 1..259
 Location/Qualifiers
 1..259
 /organism="Tubifex sp."
 /db_xref="taxon:44764"
 /dev_stage="adult"
 <1..>259
 /gene="28S rRNA"
 /product="28S ribosomal RNA"
 1..259
 /gene="28S rRNA"
 BASE COUNT 54 a 75 c 84 g 46 t
 ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 259;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTCACTGCGCGTTACTAG 22
 ||||||||||||||||||
 Db 30 CCGCTTCACTGCGCGTTACTAG 9
 RESULT 7
 TSU75861/c 276 bp RNA linear INV 23-DEC-1996
 LOCUS TSU75861
 DEFINITION Theodoxus sp. 28S ribosomal RNA gene, partial sequence.
 ACCESSION U75861
 VERSION U75861.1 GI:1750151
 KEYWORDS
 SOURCE Theodoxus sp.
 ORGANISM Theodoxus sp.
 Eukaryota; Metazoa; Mollusca; Gastropoda; Neritimorpha; Neritimorpha;
 Theodoxus.
 REFERENCE 1 (bases 1 to 276)
 AUTHORS McArthur, A.G.
 TITLE Molecular investigation of the evolutionary origins of hydrothermal vent gastropods
 JOURNAL Thesis, University of Victoria, Canada. (1996)
 REFERENCE 2 (bases 1 to 276)
 AUTHORS McArthur, A.G.
 TITLE Direct Submission
 JOURNAL Submitted (23-OCT-1996) Biology, University of Victoria, P.O. Box 3020, Victoria, BC V8W 3N5, Canada
 FEATURES
 SOURCE 1..276
 /organism="Theodoxus sp."
 /db_xref="taxon:55034"
 /dev_stage="adult"
 <1..>276
 /product="28S ribosomal RNA gene"
 BASE COUNT 61 a 76 c 91 g 48 t
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 276;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTCACTGCGCGTTACTAG 22
 ||||||||||||||||||
 Db 61 CCGCTTCACTGCGCGTTACTAG 40
 RESULT 8
 AF318260/c 283 bp DNA linear PLN 06-MAR-2001
 LOCUS AF318260
 DEFINITION Uncultured basidiomycete Cons3 24S large subunit ribosomal RNA gene, partial sequence.
 ACCESSION AF318260
 VERSION AF318260.1 GI:13236258
 KEYWORDS
 SOURCE uncultured basidiomycete Cons3.
 ORGANISM uncultured basidiomycete Cons3
 Eukaryota; Fungi; Basidiomycota; environmental samples.
 REFERENCE 1 (bases 1 to 283)
 AUTHORS Guillou, L., Nezan, E., Cuff, V., Erard-Le Denn, E., Cambon, M.-A., Gentien, P. and Barbier, G.
 TITLE Semi-nested PCR detection of three toxic dinoflagellate genera (Alexandrium, Dinophysis, and Gymnodinium) in sea water column and sediment from French coasts
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 283)
 AUTHORS Guillou, L., Nezan, E., Cuff, V., Erard-Le Denn, E., Cambon, M.-A., Gentien, P. and Barbier, G.
 TITLE Direct Submission
 JOURNAL Submitted (02-NOV-2000) DRV VP CMM, IFREMER, B.P. 70, Brest 29200,

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 218;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||

Db 59 CCGCTTCACTGCGCGTTACTAG 38

RESULT 2
E64938/c
LOCUS
DEFINITION
CDNA sequence of gene participating in induction of resistivity in
plant.
ACCESSION
E64938
VERSION
E64938.1 GI:13017969
KEYWORDS
JP 1999318477-A/70.
SOURCE
Nicotiana tabacum
ORGANISM
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanales; Solanales; Nicotiana.
REFERENCE
1 (bases 1 to 218)
AUTHORS
Peter, H.S.C.C. and Gregor, I.S.K.K.
TITLE
CDNA sequence of gene participating in induction of resistivity in
JOURNAL
Bayer AG
Parent: JP 1999318477-A 70 24-NOV-1999;

COMMENT
OS Nicotiana tabacum
PN JP 1999318477-A/70
PD 24-NOV-1999
PF 19-NOV-1999 JP 1999075762
PR 25-MAR-1998 DE 19813048.1
PI PERMER HELMUTO SCHREIER, CHRISTINE GARBERS, GREGOR LANGEN, PI
SIEGRIED KIDOROUSKI
PC C12N15/09, A01H5/00, A01N65/02, C07K14/415, C07K16/16, C12N1/15, PC
C12N1/21,
PC C12N5/10, C12N5/10, C12P21/02, C12Q1/68, G01N33/48, G01N33/56// PC
(C12N1/21, C12R1:01), (C12P21/02, C12R1:91), C12N15/00, C12N5/00, PC
C12N5/00

FEATURES
source 1..218 Location/Qualifiers
FT source 1..218
CC Key 1..218
FH Key 1..218
FT Location/Qualifiers
1..218
/organism="Nicotiana tabacum"
/db_xref="taxon:4097"
BASE COUNT 68 a 43 c 65 g 42 t
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 218;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||

Db 59 CCGCTTCACTGCGCGTTACTAG 38

RESULT 3
CS069678/c
LOCUS
DEFINITION
Cerianthus sp. 28S ribosomal RNA gene, partial sequence.
ACCESSION
U69678
VERSION
U69678.1 GI:1617542
KEYWORDS
Cerianthus sp.
SOURCE
Cerianthus sp.
ORGANISM
Eukaryota; Metazoa; Chnidaria; Anthozoa; Ceriantipatharia;
Ceriantipatharia; Ceriantipatharia; Ceriantipatharia;
REFERENCE
1 (bases 1 to 219)

AUTHORS
Chen, C.A., Odorico, D.M., ten-Lohuis, M., Veron, J.E. and Miller, D.J.
TITLE
Systematic relationships within the Anthozoa (Chnidaria: Anthozoa)
using the 5'-end of the 28S rDNA
JOURNAL
Mol. Phylogenet. Evol. 4 (2), 175-183 (1995)
MEDLINE
95392827
REFERENCE
2 (bases 1 to 219)
AUTHORS
Chen, C.A., Odorico, D.M., ten-Lohuis, M., Veron, J.E. and Miller, D.J.
TITLE
Direct Submission
JOURNAL
Submitted (06-SEP-1996) Biochemistry and Molecular Biology, James
Cook University, Townsville, Queensland 4811, Australia
Location/Qualifiers
1..219
source
/organism="Cerianthus sp."
/db_xref="taxon:51771"
rRNA
1..219
/product="28S ribosomal RNA"
BASE COUNT 61 a 48 c 59 g 51 t
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 219;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||

Db 58 CCGCTTCACTGCGCGTTACTAG 37

RESULT 4
CAU69690/c
LOCUS
DEFINITION
Corynactis australis 28S ribosomal RNA gene, partial sequence.
ACCESSION
U69690
VERSION
U69690.1 GI:1617541
KEYWORDS
Corynactis australis.
SOURCE
Corynactis australis.
ORGANISM
Corynactis australis.
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Corallimorphidae; Corynactis.
REFERENCE
1 (bases 1 to 220)
AUTHORS
Chen, C.A., Odorico, D.M., ten-Lohuis, M., Veron, J.E. and Miller, D.J.
TITLE
Systematic relationships within the Anthozoa (Chnidaria: Anthozoa)
using the 5'-end of the 28S rDNA
JOURNAL
Mol. Phylogenet. Evol. 4 (2), 175-183 (1995)
MEDLINE
95392827
REFERENCE
2 (bases 1 to 220)
AUTHORS
Chen, C.A., Odorico, D.M., ten-Lohuis, M., Veron, J.E. and Miller, D.J.
TITLE
Direct Submission
JOURNAL
Submitted (06-SEP-1996) Biochemistry and Molecular Biology, James
Cook University, Townsville, Queensland 4811, Australia
Location/Qualifiers
1..220
source
/organism="Corynactis australis"
/db_xref="taxon:51772"
rRNA
1..220
/product="28S ribosomal RNA"
BASE COUNT 54 a 55 c 67 g 44 t
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 220;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||

Db 58 CCGCTTCACTGCGCGTTACTAG 37

RESULT 5
SH065520/c
LOCUS
DEFINITION
Seriatopora hystrix 28S ribosomal RNA gene, partial sequence.
ACCESSION
U65520

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 08:07:35 : Search time 935.733 seconds

(without alignments)
492.003 Million cell updates/sec

Title: US-09-780-113D-8

Perfect score: 22

Sequence: 1 ccgcctcactgcgcgttactag 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenBank:
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_inv:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description

c	1	22	100.0	218	6	AX014514	AX014514 Sequence
c	2	22	100.0	218	6	E64938	E64938 cDNA sequen
c	3	22	100.0	219	3	CS069678	CS069678 Cerianthus
c	4	22	100.0	220	3	CA069690	CA069690 Corynactis
c	5	22	100.0	222	3	SH065520	SH065520 Seriatopora
c	6	22	100.0	223	3	TS285259	TS285259 Tuberiflex sp.
c	7	22	100.0	276	3	TS075861	TS075861 Theodoxus s
c	8	22	100.0	283	8	AF318260	AF318260 Unculture
c	9	22	100.0	285	3	AF339505	AF339505 Aplysia d
c	10	22	100.0	303	3	AF265622	AF265622 Madriactis
c	11	22	100.0	304	3	AF265642	AF265642 Caryophyl
c	12	22	100.0	305	3	AF185171	AF185171 Polyophth
c	13	22	100.0	307	3	AF185150	AF185150 Amphigen
c	14	22	100.0	307	3	AF185154	AF185154 Glycera t
c	15	22	100.0	307	3	AF185163	AF185163 Mesochaet
c	16	22	100.0	310	3	AF185165	AF185165 Sigalion
c	17	22	100.0	310	3	AF185167	AF185167 Eurythoe
c	18	22	100.0	311	3	AF185164	AF185164 Paralepid
c	19	22	100.0	317	8	AF326083	AF326083 Phialocyp
c	20	22	100.0	323	8	AF326085	AF326085 Phialocyp
c	21	22	100.0	343	3	ORDNA28S	ORDNA28S
c	22	22	100.0	347	3	MSPDNA28	MSPDNA28
c	23	22	100.0	347	3	PADNA28R	PADNA28R
c	24	22	100.0	349	3	MPDNA28SR	MPDNA28SR
c	25	22	100.0	360	3	AF358798	AF358798
c	26	22	100.0	370	3	AF358815	AF358815
c	27	22	100.0	370	3	LO28RRN5	LO28RRN5
c	28	22	100.0	371	3	PMJ225826	PMJ225826 Petrobion
c	29	22	100.0	374	3	CCJ225827	CCJ225827 Clathrina
c	30	22	100.0	417	8	FCU61652	FCU61652 Fusarium co
c	31	22	100.0	452	8	MTU43478	MTU43478 Massaria in
c	32	22	100.0	456	8	LDU43474	LDU43474 Leptosphaer
c	33	22	100.0	457	8	CDU43481	CDU43481 Leptosphaer
c	34	22	100.0	457	8	LDU43475	LDU43475 Leptosphaer
c	35	22	100.0	457	8	LDU43473	LDU43473 Leptosphaer
c	36	22	100.0	458	8	AF300719	AF300719 Salai roo
c	37	22	100.0	458	8	OFU43472	OFU43472 Ophiobolus
c	38	22	100.0	458	8	OHU43471	OHU43471 Ophiobolus
c	39	22	100.0	459	8	PTU43477	PTU43477 Pyrenophora
c	40	22	100.0	488	8	AF158614	AF158614 Fusarium
c	41	22	100.0	492	8	AF261555	AF261555 Protuber
c	42	22	100.0	518	8	PN1276065	PN1276065 Phaeococc
c	43	22	100.0	520	8	AF300718	AF300718 Salai roo
c	44	22	100.0	521	8	CHU45828	CHU45828 Candida hal
c	45	22	100.0	521	8	CMT45830	CMT45830 Candida man

ALIGNMENTS

RESULT 1
AX014514/c
LOCUS AX014514
DEFINITION Sequence 70 from Patent EP0953640.
ACCESSION AX014514
VERSION AX014514.1 GI:10040789

SOURCE

ORGANISM common tobacco.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; easterids I; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS Kiedrowski, S.D., Gardner, C.D., Jansen, G.D. and Schreier, P.H.
TITLE Cdna sequences of genes involved in the induction of resistance in plants
JOURNAL Patent: EP 0953640-A 70 03-NOV-1999;
BAYER AG (DE)

FEATURES

source location/Qualifiers
1..218 /organism="Nicotiana tabacum"
/db_xref="taxon:4097"
BASE COUNT 68 a 43 c 65 g 42 t

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 08:07:35 ; Search time 978.267 Seconds
(without alignments)
492.003 Million cell updates/sec

Title: US-09-780-113D-15

Perfect score: 23

Sequence: 1 accacgactgagcagcaccctt 23

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.com:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vi:*
30: em.htg.hum:*
31: em.htg.in:*
32: em.htg.other:*
33: em.htgo.in:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

C 1	23	100.0	712	8	AF042820	AF042820 Heterosig
C 2	23	100.0	713	8	AF086948	AF086948 Heterosig
C 3	23	100.0	715	8	AF210741	AF210741 Heterosig
C 4	23	100.0	715	8	AF211256	AF211256 Heterosig
C 5	23	100.0	4103	8	AF409124	AF409124 Heterosig
C 6	18.2	79.1	126149	9	AC006352	AC006352 Homo sapi
C 7	18.2	79.1	169710	2	AL627202	AL627202 Homo sapi
C 8	18.2	79.1	188930	2	AL592114	AL592114 Homo sapi
C 9	18.2	79.1	197083	2	AC009106	AC009106 Mus muscu
C 10	18.2	79.1	214025	2	AC007882	AC007882 Homo sapi
C 11	18.2	79.1	234542	2	AC007908	AC007908 Homo sapi
C 12	18.2	79.1	234542	2	HUAC002041	HUAC002041 Homo sapi
C 13	17.8	77.4	215359	9	AC012495	AC012495 Homo sapi
C 14	17.4	75.7	34105	2	CNS000VE	CNS000VE Homo sapi
C 15	17.4	75.7	197764	2	AC021605	AC021605 Homo sapi
C 16	17.2	74.8	26689	9	AL499629	AL499629 Human DNA
C 17	17.2	74.8	43772	9	AC004791	AC004791 Homo sapi
C 18	17.2	74.8	46708	2	AL359434	AL359434 Human DNA
C 19	17.2	74.8	73806	2	AC095461	AC095461 Rattus no
C 20	17.2	74.8	73108	9	AC005363	AC005363 Homo sapi
C 21	17.2	74.8	80668	9	AC005606	AC005606 Homo sapi
C 22	17.2	74.8	81579	9	AE006640	AE006640 Homo sapi
C 23	17.2	74.8	107549	9	AC074178	AC074178 Homo sapi
C 24	17.2	74.8	123331	9	AC005164	AC005164 Homo sapi
C 25	17.2	74.8	128628	9	AC093758	AC093758 Homo sapi
C 26	17.2	74.8	163031	9	AC024947	AC024947 Homo sapi
C 27	17.2	74.8	163889	2	AC004958	AC004958 Homo sapi
C 28	17.2	74.8	175996	2	AC010539	AC010539 Homo sapi
C 29	17.2	74.8	184919	9	AF069291	AF069291 Homo sapi
C 30	17.2	74.8	207962	2	AC015679	AC015679 Homo sapi
C 31	17.2	74.8	238169	2	AC091396	AC091396 Mus muscu
C 32	17.2	74.8	242965	2	AC023175	AC023175 Mus muscu
C 33	17.2	74.8	320250	2	AF117829	AF117829 Homo sapi
C 34	17.2	73.9	154921	2	AC109793	AC109793 Bos tauru
C 35	17.2	73.9	205140	2	AC109795	AC109795 Bos tauru
C 36	17.2	73.9	335372	2	AC094935	AC094935 Rattus no
C 37	16.8	73.0	280	9	HSCRP06	HSCRP06 Homo sapi
C 38	16.8	73.0	355	13	AF349055	AF349055 Unculture
C 39	16.8	73.0	42856	2	AC100075	AC100075 Mus muscu
C 40	16.8	73.0	110000	2	AC105915	AC105915 Homo sapi
C 41	16.8	73.0	110000	2	AC005079_2	AC005079_2 Continuation (3 of
C 42	16.8	73.0	121459	2	AL137860	AL137860 Homo sapi
C 43	16.8	73.0	168042	2	AC022529	AC022529 Homo sapi
C 44	16.8	73.0	170128	2	AC005059	AC005059 Homo sapi
C 45	16.8	73.0	175390	2	AC094641	AC094641 Rattus no

ALIGNMENTS

RESULT 1
LOCUS AF042820/c 712 bp DNA linear PLN 01-FEB-1998
DEFINITION Heterosigma akashiwo 24S large subunit ribosomal RNA-sequence
ACCESSION AF042820
VERSION AF042820.1 GI:2827390
KEYWORDS
SOURCE
ORGANISM Heterosigma akashiwo.
Heterosigma akashiwo.
Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
REFERENCE
1 (bases 1 to 712)
Lee, S.W., Park, C.G. and Park, Y.S.
24S ribosomal RNA sequence analysis of dinoflagellates isolated from red-tide in southern coast of Korea
TITLE
Unpublished
JOURNAL
REFERENCE
1 (bases 1 to 712)
Lee, S.W., Park, C.G. and Park, Y.S.
Direct Submission
JOURNAL
Submitted (14-JAN-1998) Department of Microbiology, Inje University, Ochangdong 607, Kimae 621-749, Korea
FEATURES
source
1. 712

/organism="Heterosigma akashiwo"
/db_xref="taxon:2829"
/note="Isolated from red-tide sea water in Masan Bay, Korea; classified by morphological studies"
<1..>712
/note="variable domains D1 and D2"
/product="24S large subunit ribosomal RNA"
BASE COUNT 164 a 149 c 211 g 188 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAGAGCTGAGCAGCACCCTT 23
|||||
Db 149 ACCAGAGCTGAGCAGCACCCTT 127

RESULT 2
AF086948/c 713 bp DNA linear PLN 15-OCT-1998
LOCUS Heterosigma akashiwo large subunit ribosomal RNA gene, partial
DEFINITION
Sequence.
AF086948
VERSION AF086948.1 GI:3695263
KEYWORDS
SOURCE Heterosigma akashiwo.
ORGANISM Heterosigma akashiwo
REFERENCE 1 (bases 1 to 713)
AUTHORS Connell, L.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) Environmental Conservation Division,
National Marine Fisheries Service, 2725 Montlake Blvd East,
Seattle, WA 98112, USA
FEATURES
source 1..713
Location/Qualifiers
1..713
/organism="Heterosigma akashiwo"
/strain="CCMP-452"
/db_xref="taxon:2829"
<1..>713
/note="contains variable domains D1-D3"
/product="large subunit ribosomal RNA"
BASE COUNT 165 a 148 c 210 g 190 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 713;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACCAGAGCTGAGCAGCACCCTT 23
|||||
Db 147 ACCAGAGCTGAGCAGCACCCTT 125

RESULT 3
AF210741/c 715 bp DNA linear "N 02-JAN-2001
LOCUS Heterosigma akashiwo large subunit ribosomal RNA gene, partial
DEFINITION
Sequence.
AF210741
VERSION AF210741.1 GI:12003313
KEYWORDS
SOURCE Heterosigma akashiwo.
ORGANISM Heterosigma akashiwo
REFERENCE 1 (bases 1 to 715)
AUTHORS Tyrell, J.V., Scholm, C.A., Bergquist, P.R., and Bergquist, P.L.
TITLE Detection and Enumeration of Heterosigma akashiwo and Fibrocapsa
japonica (Raphidophyceae) using rRNA-targeted Oligonucleotide
Probes
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 715)
AUTHORS Tyrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Phylogeny of the Raphidophytes Based on Large-Subunit rRNA Gene
Sequences
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 715)
AUTHORS Tyrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1999) Research and Development, Monterey Bay
Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA
95039-0628, USA
FEATURES
source Location/Qualifiers
1..715
/organism="Heterosigma sp. CAMR10"
/strain="CAMR10"
/db_xref="taxon:147349"
/note="similar to Heterosigma akashiwo sequence AF210741"
<1..>715
/note="D1 and D2 domains; similar to Heterosigma akashiwo
sequence presented in Genbank Accession Number AF210741"
/product="large subunit ribosomal RNA"
BASE COUNT 166 a 148 c 212 g 189 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 715;
Best Local Similarity 100.0%; Pred. No. 0.15;

OY 1 ACCAGAGCTGAGCAGCACCCTT 23
|||||
Db 149 ACCAGAGCTGAGCAGCACCCTT 127

RESULT 4
AF211256/c 715 bp DNA linear PLN 11-JAN-2001
LOCUS Heterosigma sp. CAMR10 large subunit ribosomal RNA gene, partial
DEFINITION
Sequence.
AF211256
VERSION AF211256.1 GI:12082493
KEYWORDS
SOURCE Heterosigma sp. CAMR10.
ORGANISM Heterosigma sp. CAMR10
REFERENCE 1 (bases 1 to 715)
AUTHORS Tyrell, J.V., Scholm, C.A., Bergquist, P.R., and Bergquist, P.L.
TITLE Detection and Enumeration of Heterosigma akashiwo and Fibrocapsa
japonica (Raphidophyceae) Using rRNA-targeted Oligonucleotide
Sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 715)
AUTHORS Tyrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Phylogeny of the Raphidophytes Based on Large-Subunit rRNA Gene
Sequences
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 715)
AUTHORS Tyrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1999) Research and Development, Monterey Bay
Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA
95039-0628, USA
FEATURES
source Location/Qualifiers
1..715
/organism="Heterosigma sp. CAMR10"
/strain="CAMR10"
/db_xref="taxon:147349"
/note="similar to Heterosigma akashiwo sequence AF210741"
<1..>715
/note="D1 and D2 domains; similar to Heterosigma akashiwo
sequence presented in Genbank Accession Number AF210741"
/product="large subunit ribosomal RNA"
BASE COUNT 166 a 148 c 212 g 189 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 715;
Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 23: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ACCAGACTGAGCAGCACCCTT 23
 |||||||
 Db 149 ACCAGACTGAGCAGCACCCTT 127

RESULT 5
 AF409124/c 4103 bp DNA linear PLN 23-OCT-2001
 LOCUS AF409124
 DEFINITION Heterosigma akashiwo internal transcribed spacer 1, 5.8S ribosomal
 RNA gene, internal transcribed spacer 2, and 28S large subunit
 ribosomal RNA gene, complete sequence.
 ACCESSION AF409124 GI:16326556
 VERSION AF409124.1
 KEYWORDS Heterosigma akashiwo.
 SOURCE Heterosigma akashiwo
 ORGANISM Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.

REFERENCE 1 (bases 1 to 4103)
 Ben Ali, A., De Baere, R., De Wachter, R. and Van de Peer, Y.
 Evolutionary relationships among heterokont algae (the autotrophic
 stramenopiles) based on a combined analysis of small and large
 subunit RNA

JOURNAL unpublished
 REFERENCE 2 (bases 1 to 4103)
 Ben Ali, A.
 AUTHORS Direct Submission
 TITLE Submitted (14-AUG-2001) Biochemistry, University of Antwerp,
 Universiteitsplein 1, Wilrijk 2610, Belgium
 JOURNAL location/Qualifiers

FEATURES
 source 1..4103
 /organism="Heterosigma akashiwo"
 /db_xref="taxon:2829"
 misc_RNA 1..289
 /product="Internal transcribed spacer 1"
 rRNA 290..447
 /product="5.8S ribosomal RNA"
 misc_RNA 448..650
 /product="Internal transcribed spacer 2"
 rRNA 651..4103
 /product="28S large subunit ribosomal RNA"

BASE COUNT 1045 a 899 c 1142 g 1017 t

ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 4103;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGACTGAGCAGCACCCTT 23
 |||||||
 Db 824 ACCAGACTGAGCAGCACCCTT 802

RESULT 6
 AC006352/c 126149 bp DNA linear PRI 30-SEP-2000
 LOCUS AC006352
 DEFINITION Homo sapiens PAC clone RP5-1044J9 from 7p14-p12, complete sequence.
 ACCESSION AC006352
 VERSION AC006352.3 GI:4753268
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 126149)
 Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 REFERENCE 2 (bases 1 to 126149)
 Du, F., Wohlmann, P. and McGrane, B.
 AUTHORS The sequence of Homo sapiens PAC clone RP5-1044J9
 TITLE

JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 126149)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (11-JAN-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 4 (bases 1 to 126149)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAY-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 5 (bases 1 to 126149)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On May 5, 1999 this sequence version replaced gi:4337277.

COMMENT ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@watsn.wustl.edu
 ----- Summary Statistics
 Center project name: H.DJ1044J09

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
 http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
 mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 This clone was derived from human PAC library RPCI-5, prepared by
 Pieter de Jong and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu) using the method described by
 Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from
 one male donor.
 The clone may be obtained either from Genome Systems, Inc.
 (http://www.genomesystems.com) or Research Genetics, Inc.
 (http://www.resgen.com); or from Pieter de Jong.
 VECTOR: pCYFAC2

NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP5-1044J9;
 actual end is at 126149 of RP5-1044J9.

FEATURES
 source 1..126149
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7p14-p12"
 /clone="RP5-1044J9"
 /clone_id="RPCI-5"
 misc_feature 323..411

```

misc_feature /note="similar to EST W29670 (NID:g611768)"
337..411
/note="similar to EST AA569186 (NID:g2342240) nm30d10.s1"
misc_feature 494..795
/note="similar to EST AA569186 (NID:g2342240) nm30d10.s1"
misc_feature 575..1056
/note="similar to EST AA948436 (NID:g3109689) on67a01.s1"
repeat_region 1035..1468
/rpt_family="L1"
repeat_region 1590..1602
/rpt_family="(TA)n"
repeat_region 1603..1650
/rpt_family="(CA)n"
repeat_region 1651..1661
/rpt_family="(TA)n"
repeat_region 1798..2689
/rpt_family="L1"
repeat_region 2657..2824
/rpt_family="L1"
repeat_region 6125..6206
/rpt_family="MER4-group"
repeat_region 6601..7454
/rpt_family="MER4-group"
repeat_region 7455..7520
/rpt_family="Alu"
repeat_region 7521..8553
/rpt_family="MER4-group"
repeat_region 8554..8830
/rpt_family="Alu"
repeat_region 8831..11043
/rpt_family="MER4-group"
repeat_region 11044..11411
/rpt_family="MAlR"
repeat_region 11412..11453
/rpt_family="MER4-group"
repeat_region 11455..11517
/rpt_family="MER4-group"
repeat_region 11522..11627
/rpt_family="Retroviral"
repeat_region 11614..11909
/rpt_family="LTR198"
repeat_region 11829..12044
/rpt_family="Retroviral"
repeat_region 12045..12268
/rpt_family="MER4-group"
repeat_region 12284..12682
/rpt_family="MER4-group"
repeat_region 12683..12924
/rpt_family="MAlR"
repeat_region 12921..13193
/rpt_family="MER4-group"
repeat_region 14444..14865
/rpt_family="MER4-group"
repeat_region 14961..15275
/rpt_family="Alu"
repeat_region 15394..15993
/rpt_family="MER4-group"
repeat_region 16088..16116
/rpt_family="AT-rich"
repeat_region 16142..16589
/rpt_family="L1"
misc_feature 17241..17338
/note="similar to Mus musculus EST AA170256 (NID:g1748794)
ms87g10.r1"
misc_feature 17241..17338
/note="similar to Mus musculus EST AA575218 (NID:g2349844)
vhi9a01.r1"
misc_feature 17266..17338
/note="similar to EST R72787 (NID:g846819) yj91h09.r1"
misc_feature 17447..17748
/note="similar to EST R72787 (NID:g846819) yj91h09.r1"
misc_feature 17447..17750
/note="similar to Mus musculus EST AA170256 (NID:g1748794)

```

```

misc_feature ms87g10.r1"
17447..17751
/note="similar to Mus musculus EST AA575218 (NID:g2349844)
vhi9a01.r1"
misc_feature 17490..17780
/note="similar to EST H73816 (NID:g1046750) ys11b01.r1"
misc_feature 17492..17750
/note="similar to EST AA988730 (NID:g3174301) or79h12.s1"
repeat_region 17978..18004
/rpt_family="AT-rich"
repeat_region 18694..19007
/rpt_family="MER4-group"
repeat_region 20166..20452
/rpt_family="L1"
repeat_region 20456..20498
/rpt_family="L1"
repeat_region 22114..22213
/rpt_family="L1"
misc_feature 22502..22769
/note="match to EST AA489905 (NID:g220780) ab05c02.r1"
repeat_region 22573..22704
/rpt_family="L1"
misc_feature 23048..23453
/note="similar to EST N52241 (NID:g193375) yv46q12.s1"
misc_feature 24524..24597
/note="similar to EST AA300571 (NID:g1953132)"
misc_feature 24682..24840
/note="similar to EST AA300571 (NID:g1953132)"
misc_feature 24682..24831
/note="similar to EST T29670 (NID:g611768)"
repeat_region 24954..25460
/rpt_family="L1"
repeat_region 25461..25486
/rpt_family="(CA)n"
repeat_region 25487..25557
/rpt_family="L1"
repeat_region 25550..26548

```

```

Query Match 79.1%; Score 18.2; DB 9; Length 126149;
Best Local Similarity 87.0%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCAGCAGCAGCAGCAGCAGCCTT 23
Db 123610 ACCAGCAGCAGCAGCAGCAGCCTT 123588
|||||
RESULT 7
AL627202/c 169710 bp DNA linear HTG 07-NOV-2001
LOCUS Homo sapiens chromosome 1 clone Rpl1-174E23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 12 unordered pieces.
ACCESSION AL627202
VERSION HTG: HTGS_PHASE1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Clark, S.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hunger@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 30, 2001 this sequence version replaced g1:16501251.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hunger@sanger.ac.uk
----- Project Information
Center project name: BA174P23

```

```

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 165366 bases at least Q40
Consensus quality: 167987 bases at least Q30
Insert size: 168610; sum-of-contigs
Insert size: 172230; 12.6% error; agarose-fp
Quality coverage: 5.03x in Q20 bases; sum-of-contigs quality
coverage: 5.15x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 28829: contig of 28829 bp in length
28830 28929: gap of 100 bp
28930 35872: contig of 6943 bp in length
35873 35972: gap of 100 bp
35973 48765: contig of 12793 bp in length
48766 48865: gap of 100 bp
48866 77651: contig of 28786 bp in length
77652 77751: gap of 100 bp
77752 83328: contig of 5577 bp in length
83329 83428: gap of 100 bp
83429 89376: contig of 5948 bp in length
89377 89476: gap of 100 bp
89477 95053: contig of 5577 bp in length
95054 95153: gap of 100 bp
95154 115825: contig of 20672 bp in length
115826 115925: gap of 100 bp
115926 117990: contig of 2065 bp in length
117991 118090: gap of 100 bp
118091 129051: contig of 10961 bp in length
129052 129151: gap of 100 bp
129152 155347: contig of 26196 bp in length
155348 155447: gap of 100 bp
155448 169710: contig of 14263 bp in length.
Location/Qualifiers

1 169710
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-174P23"
/clone.lib="RPC1-11.1"
1. 28829
/note="assembly_fragment:00723
fragment_chain:1"
misc_feature
28930..35872
/note="assembly_fragment:03428
fragment_chain:1"
misc_feature
35973..48765
/note="assembly_fragment:02971
fragment_chain:1"
misc_feature
48866..77651
/note="assembly_fragment:02451
fragment_chain:1"
misc_feature
77752..83328
/note="assembly_fragment:02530
fragment_chain:1"
misc_feature
83429..89376
/note="assembly_fragment:01668
fragment_chain:1"
misc_feature
89477..95053
/note="assembly_fragment:02114
fragment_chain:2"
misc_feature
95154..115825
/note="assembly_fragment:03342

```

```

fragment_chain:2"
115926..117990
/note="assembly_fragment:03173
fragment_chain:2"
misc_feature
118091..129051
/note="assembly_fragment:01101
fragment_chain:2"
misc_feature
129152..155347
/note="assembly_fragment:02969"
155448..169710
/note="assembly_fragment:01462
clone_end:17
vector_side:right"
BASE COUNT 44171 a 42638 c 39303 g 42488 t 1110 others
ORIGIN
Query Match 79.1%; Score 18.2; DB 2; Length 169710;
Best Local Similarity 87.0%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCAGCACTGACGACGACCTT 23
||||| ||||| ||||| ||
Db 80661 ACCAGCACTGAGGACGACCTT 80639

RESULT 8
AL592114 188930 bp DNA linear HTG 03-JAN-2002
LOCUS
AL592114
DEFINITION
Homo sapiens chromosome 1 clone RP11-203P10, *** SEQUENCING IN
PROGRAMS ***, 2 unordered pieces.
AL592114
VERSION
AL592114.9 GI:18070892
KEYWORDS
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS
Clark,S.
TITLE
Direct Submission
JOURNAL
Submitted (02-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk clone requests: clonequest@sanger.ac.uk
On Jan 4, 2002 this sequence version replaced gi:16501196.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: ba203P10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M77815; 1% of reads
Sequencing vector: plasmid; 108752; 98% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 186643 bases at least Q40
Consensus quality: 188754 bases at least Q30
Consensus quality: 188773 bases at least Q20
Insert size: 188830; sum-of-contigs
Insert size: 187990; 6.8% error; agarose-fp
Quality coverage: 11.94x in Q20 bases; sum-of-contigs quality
coverage: 12.06x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 34047: contig of 34047 bp in length

```


RESULT	10
LOCUS	AC007882 214025 bp DNA linear PRI 30-SEP-2000
DEFINITION	Homo sapiens BAC clone RP11-499D5 from gp11.2-q11.2, complete sequence.
ACCESSION	AC007882
VERSION	AC007882.3 GI:7625250
KEYWORDS	HTS.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 214025)
AUTHORS	Tolson,J.E. and Waterston,R.
TITLE	Toward a complete human genome sequence
JOURNAL	Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE	99063792
REFERENCE	2 (bases 1 to 214025)
AUTHORS	Reltz,L., Stoneking,T. and Andrews,S.
TITLE	The sequence of Homo sapiens BAC clone RP11-499D5
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 214025)
TITLE	Waterston,R.H.
AUTHORS	Direct Submission
JOURNAL	Submitted (19-JUN-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	4 (bases 1 to 214025)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (20-APR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 214025)
AUTHORS	Wa erston,R.H.
TITLE	I..ect Submission
JOURNAL	Submitted (21-APR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	6 (bases 1 to 214025)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (22-APR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	7 (bases 1 to 214025)
AUTHORS	Waterston,R.
TITLE	Direct Submission
JOURNAL	Submitted (20-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Apr 20, 2000 this sequence version replaced gi:5636167.
COMMENT	-- Genome Center Center: Washington Unlversity Genome Sequencing Center Center code: WUGSC Web site: http://genome.wustl.edu/gsc Contact: saplense@watson.wustl.edu -- Summary Statistics Center project name: H_NH0499D05 -- NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions. This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

AC007908
 VERSION AC007908.3 GI:9094205
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 227856)
 DOE Joint Genome Institute.
 TITLE
 JOURNAL Sequencing of Human Chromosome 16
 REFERENCE
 2 (bases 1 to 227856)
 Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Mehncke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurry, K., Han, C. and Deaven, L.
 TITLE
 JOURNAL Direct Submission
 Submitted (24-JUN-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 On Jul 13, 2000 this sequence version replaced gi:7211873.
 COMMENT
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

 Project Information
 Center Project Name: 0
 Center clone name: RPci-11_49905

 Summary Statistics
 Consensus quality: 184734 bases at least Q40
 Consensus quality: 201642 bases at least Q30
 Consensus quality: 209314 bases at least Q20
 Estimated insert size: 218590; agarose-fp estimation
 Estimated insert size: 224956; sum-of-contigs estimation
 Quality coverage: 4.43 in Q20 bases; agarose-fp estimation
 Quality coverage: 4.3 in Q20 bases; sum-of-contigs estimation
 NOTE: This is a 'working draft' sequence. It currently
 consists of 30 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 1259: contig of 1259 bp in length
 * 1260 1359: gap of unknown length
 * 1360 2455: contig of 1096 bp in length
 * 2456 2555: gap of unknown length
 * 2556 3691: contig of 1136 bp in length
 * 3692 3791: gap of unknown length
 * 3792 4805: contig of 1014 bp in length
 * 4806 4905: gap of unknown length
 * 4906 5938: contig of 1033 bp in length
 * 5939 6038: gap of unknown length
 * 6039 7051: contig of 1013 bp in length
 * 7052 7151: gap of unknown length
 * 7152 8247: contig of 1095 bp in length
 * 8248 8347: gap of unknown length
 * 8348 9994: contig of 1647 bp in length
 * 9995 10093: gap of unknown length
 * 10094 11375: contig of 1282 bp in length
 * 11376 11475: gap of unknown length
 * 11476 12513: contig of 1038 bp in length
 * 12514 12613: gap of unknown length
 * 12614 13645: contig of 1032 bp in length
 * 13646 13745: gap of unknown length
 * 13746 15252: contig of 1507 bp in length
 * 15253 15353: gap of unknown length
 * 15354 16573: contig of 1221 bp in length
 * 16574 16773: gap of unknown length

16674 18205: contig of 1532 bp in length
 * 18206 18305: gap of unknown length
 * 18306 19913: contig of 1608 bp in length
 * 19914 20013: gap of unknown length
 * 20014 22479: gap of unknown length
 * 22480 22579: gap of unknown length
 * 22580 25293: contig of 2714 bp in length
 * 25294 25393: gap of unknown length
 * 25394 29118: contig of 3725 bp in length
 * 29119 29219: gap of unknown length
 * 29220 32190: contig of 2971 bp in length
 * 32191 32290: gap of unknown length
 * 32291 33294: contig of 1005 bp in length
 * 33295 33393: gap of unknown length
 * 33394 34509: contig of 1115 bp in length
 * 34510 34610: gap of unknown length
 * 34611 36147: contig of 1558 bp in length
 * 36148 36247: gap of unknown length
 * 36248 42248: contig of 6001 bp in length
 * 42249 42348: gap of unknown length
 * 42349 52325: contig of 9977 bp in length
 * 52326 52425: gap of unknown length
 * 52426 60479: contig of 8054 bp in length
 * 60480 60579: gap of unknown length
 * 60580 68636: contig of 8057 bp in length
 * 68637 68737: gap of unknown length
 * 68738 83105: contig of 14369 bp in length
 * 83106 83205: gap of unknown length
 * 83206 116897: contig of 33692 bp in length
 * 116898 116997: gap of unknown length
 * 116998 165561: contig of 48564 bp in length
 * 165562 165661: gap of unknown length
 * 165662 227856: contig of 62195 bp in length.
 Location/Qualifiers
 1..227856
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-499D5"
 /clone_1bp="RPci human BAC library 11"
 BASE COUNT 68439 a 46078 c 44943 g 64859 t 3537 others
 ORIGIN
 Query Match 79.1%; Score 18.2; DB 2; Length 227856;
 Best Local Similarity 87.0%; Pred. No. 69;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ACCGACGACGAGCAGCAGCTTT 23
 ||||| ||| ||||| ||||| |||
 DB 130616 ACCGAGCAGCAGCAGCAGCTT 130594
 RESULT 12
 HUC002041 234542 bp DNA linear PRI 23-NOV-1999
 LOCUS Human Chromosome 16 BAC clone C179875K-A-11E1, complete sequence.
 AC002041
 VERSION AC002041.1 GI:2576343
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 234542)
 Loftus, B.J., Kim, Y.J., Shedd, V.P., Kalush, F., Brandon, R.,
 Fuhrmann, J., Mason, T., Crosby, M.L., Barstead, M., Cronin, L.,
 Deslattes Mays, A., Cao, Y., Xu, R.X., Kang, H.L., Mitchell, S.,
 Eichler, E.E., Harris, P.C., Venter, J.C. and Adams, M.D.
 TITLE
 JOURNAL Genome duplications and other features in 12 Mb of DNA sequence
 from human chromosome 16p and 16q
 GENOMICS 99425270
 MEDLINE
 PUBMED 10493829

REFERENCE AUTHORS	2 (bases 1 to 234542) Adams,M.D., Loftus,B.J., Zhou,L., Crosby,M., Fuhrmann,J., Brandon,R., Kim,U.J., Kertavagae,A.R. and Venter,J.C. Human Chromosome 16 BAC clone C179678K-A-17E1
JOURNAL TITLE	Unpublished 3 (bases 1 to 234542)
REFERENCE AUTHORS	Adams,M.D. and Loftus,B.J. Submitted Submission Direct Submission Submitted (29-APR-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
JOURNAL TITLE	4 (bases 1 to 234542) Adams,M.D. Direct Submission Submitted (30-OCT-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
REFERENCE AUTHORS	5 (bases 1 to 234542) Adams,M.D. Direct Submission Submitted (31-DEC-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
JOURNAL TITLE	6 (bases 1 to 234542) Adams,M.D. and Loftus,B.J. Direct Submission Submitted (17-MAR-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
REFERENCE AUTHORS	7 (bases 1 to 234542) Adams,M.D. Direct Submission Submitted (11-SEP-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
JOURNAL TITLE	On Oct 30, 1997 this sequence version replaced q1:2570252. Address all correspondence to: Mark Adams The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT	e-mail address: mdamas@tigr.org. The bac location is on chromosome BAC clone is located on human chromosome 16p11.2. The orientation of the sequence is from Sp6 to T7 end. Genes were identified by a combination of five methods including: XGRAIL (available by anonymous ftp from athur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), GenScan (Chris Burge, http://genome.stanford.edu/~chris/GENSCANW.html) searches of the complete sequence against a peptide database, and the human gene index database at TIGR (http://www.tigr.org/tigr/hgi/hgi.html). A gene with homology to another protein is annotated as the isolog of that protein. Genes without peptide homology having spliced EST hits are termed 'unknown protein'. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES	Location/Qualifiers
SOURCE	1.234542 "Organism=Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16p11.2" /clone="A-17E1" 36147..39310 /gene="A-17E1.9" /pseudo 59641..87064 /gene="A-17E1.1" /note="pseudogene" /pseudo 87762..87893 /note="32826, A002D07, Chr. -, Homo sapiens" /db_xref="dbSTS:G19948" complement(128248..129149) /gene="A-17E1.2" /note="pseudogene" /pseudo 132476..132730 /note="42645, SWSS4047, Chr. 7, Homo sapiens" /db_xref="dbSTS:G28745"
STS	

STS	136754	136924	/note="42758, D11S1053, Chr. -, Homo sapiens"
			/db_xref="dbSTS:c28812"
STS	137107	137238	/note="32826, A002D07, Chr. -, Homo sapiens"
			/db_xref="dbSTS:G19948"
STS	147776	149953	/note="7821, STS1-cSRU-31b6-uA/cSRU-31b6-uZ, Chr. -, Homo sapiens"
			/db_xref="dbSTS:G02335"
gene	159399	159940	/gene="A-17E1.8"
			/pseudo
STS	160622	160753	/note="32826, A002D07, Chr. -, Homo sapiens"
			/db_xref="dbSTS:G19948"
STS	163896	164027	/note="32826, A002D07, Chr. -, Homo sapiens"
			/db_xref="dbSTS:G19948"
			complement(169812..202126)
			/note="creatine transporter and CDM gene containing duplication arising from Chx428"
gene	181687	183736	/gene="A-17E1.4"
			/pseudo
STS	190412	190532	/note="18040, SHCG-9731, Chr. 16, Homo sapiens"
			/db_xref="dbSTS:G11301"
gene	complement(190561..197183)		/gene="A-17E1.5"
			/pseudo
gene	206967	207475	/gene="A-17E1.6"
			/pseudo
gene	216986	217560	/gene="A-17E1.10"
			/pseudo
gene	228452	229030	/gene="A-17E1.11"
			/pseudo
BASE COUNT	63242 a	50172 c	50664 g 70414 t 50 others
ORIGIN			
Query Match			79.1%: Score 18.2; DB 9; Length 234542;
Best Local Similarity			87.0%: Pred. No. 69;
Matches	20; Conservative	0; Mismatches	3; Indels 0; Gaps 0;
OY	1	ACCAACACTGAGACGACACCTT	23
Db	81049	ACCAAGTACGACGACGACCTT	81071
RESULT 13			
LOCUS	AC012495	215359 bp	DNA linear PRI 07-NOV-2001
ACCESSION	AC012495	BAC clone RP11-327f6	from 2, complete sequence.
VERSION	AC012495.9	GI:16756348	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 215359)		
AUTHORS	Sulston, J.E. and Waterston, R.		
TITLE	Toward a complete human genome sequence		
JOURNAL	Genome Res. 8 (11), 1097-1108 (1998)		
MEDLINE	99063792		
REFERENCE	2 (bases 1 to 215359)		
AUTHORS	Du, F., Haakenson, W., Elliott, G. and Richards, M.		
TITLE	The sequence of Homo sapiens BAC clone RP11-327f6		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 215359)		

AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 215359)
REFERENCE Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 7, 2001 this sequence version replaced gi:13431146.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0327F06

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatemio,M., Catanese,J.J. and de Jong,P.P. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-434B12; the clone sequenced to the right is AC051646. Actual start of this clone is at base position 1 of RP11-327F6; actual end is at base position 215359 of RP11-327F6.

The sequence H_NH0327F06 from base positions 115714 to 115728 is represented by sequence derived from PCR on BAC DNA.

FEATURES

Source

1. 215359

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-327F6"

/clone.lib="RPCI-11"

164..293

repeat_region

/rpt_family="Alu"

473..757

repeat_region

/rpt_family="Alu"

846..1142

repeat_region

/rpt_family="Alu"

2744..3066

repeat_region

/rpt_family="MER2_type"

repeat_region 4273..4576
/rpt_family="Alu"
repeat_region 8207..8227
/rpt_family="AT-rich"
repeat_region 8476..8607
/rpt_family="Alu"
repeat_region 10228..10586
/rpt_family="ERV1"
repeat_region 10727..10757
/rpt_family="AT-rich"
repeat_region 11462..11768
/rpt_family="Alu"
repeat_region 14106..14201
/rpt_family="L1"
repeat_region 14307..15160
/rpt_family="L1"
repeat_region 17498..17825
/rpt_family="Alu"
repeat_region 18341..18600
/rpt_family="Alu"
repeat_region 18942..19033
/rpt_family="L2"
repeat_region 19700..19759
/rpt_family="MIR"
repeat_region 20267..20410
/rpt_family="(TATA)n"
repeat_region 20576..20600
/rpt_family="(A)n"
repeat_region 20601..20657
/rpt_family="(TA)n"
repeat_region 21725..22018
/rpt_family="Alu"
repeat_region 22203..22249
/rpt_family="Alu"
repeat_region 22316..22616
/rpt_family="Alu"
repeat_region 22657..22701
/rpt_family="(T)n"
repeat_region 22731..22752
/rpt_family="(TAAA)n"
repeat_region 24055..24075
/rpt_family="AT-rich"
repeat_region 24507..24559
/rpt_family="MIR"
repeat_region 27464..27498
/rpt_family="(T)n"
repeat_region 27985..28215
/rpt_family="Alu"
repeat_region 28243..28526
/rpt_family="Alu"
repeat_region 28529..28552
/rpt_family="AT-rich"
repeat_region 28752..28843
/rpt_family="MIR"
repeat_region 29211..29265
/rpt_family="Mariner"
repeat_region 29352..29615
/rpt_family="L1"
repeat_region 29647..29795
/rpt_family="L1"
repeat_region 30200..30220
/rpt_family="AT-rich"
repeat_region 31406..31724
/rpt_family="L2"
repeat_region 31839..32706
/rpt_family="L1"
repeat_region 32707..33079
/rpt_family="MIR"
repeat_region 33080..33160
/rpt_family="L1"
repeat_region 34069..34373
/rpt_family="Alu"
repeat_region 34685..34962

```

repeat_region /fpl_family="L1"
36033..36204 /fpl_family="MIR"
repeat_region /fpl_family="MIR"
38668..38979 /fpl_family="Alu"
repeat_region /fpl_family="Alu"
39472..39493 /fpl_family="MIR"
repeat_region /fpl_family="MIR"
39494..39702 /fpl_family="MER2_type"
repeat_region /fpl_family="MIR"
39703..39785 /fpl_family="MIR"
repeat_region /fpl_family="L2"
39794..39843 /fpl_family="L2"
repeat_region /fpl_family="L2"
40006..40612 /fpl_family="L1"
repeat_region /fpl_family="L2"
40797..40871 /fpl_family="L2"
repeat_region /fpl_family="L2"
41239..41319 /fpl_family="L2"
repeat_region /fpl_family="L2"
41909..42079 /fpl_family="Alu"
repeat_region /fpl_family="Alu"
42255..42861 /fpl_family="L2"
repeat_region /fpl_family="L2"
43465..43500 /fpl_family="AT_rich"
repeat_region /fpl_family="AT_rich"
45073..45094 /fpl_family="AT_rich"
repeat_region /fpl_family="AT_rich"
46334..46624 /fpl_family="Alu"
repeat_region /fpl_family="AT_rich"
46625..46654 /fpl_family="AT_rich"
repeat_region /fpl_family="AT_rich"
47891..47941 /fpl_family="AT_rich"
repeat_region /fpl_family="AT_rich"
48406..48703 /fpl_family="Alu"
repeat_region /fpl_family="Alu"

Query Match 77.4% Score 17.8; DB 9; Length 215359;
Best Local Similarity 90.5%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCT 21
|||||
Db 141309 ACCACGACTGACGACCT 141329

RESULT 14
CNS00YVE 34105 bp DNA linear PRI 11-OCT-1999
LOCUS Homo sapiens genome region containing hypervariable minisatellites
DEFINITION chromosome 8[8q24.3] of Homo sapiens.
ACCESSION AL0096807.1 GI:5441561
VERSION 1
KEYWORDS hypervariable; minisatellite.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 34105)
AUTHORS Petit,E., Vico,V., Brothier,P. and Vergnaud,G.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-1999) Laboratoire IECH - Institut de Genetique et
Microbiologie Universite Paris XI 91400 ORSAY - FRANCE & Genoscope
& Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE
& UGM - Centre d'Etudes du Bouchet BP3 91710 Verri le Petit - FRANCE
2 (bases 1 to 34105)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (05-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
FEATURES
source 1..34105
Location/Qualifiers
1..34105
/organism="Homo sapiens"

```

```

/dp_xref="taxon:9606"
/chromosome="8"
/clone="CEBA2"
/map="8q24.3"

BASE COUNT 8661 a 8558 c 8659 g 8227 t
ORIGIN

Query Match 75.7% Score 17.4; DB 9; Length 34105;
Best Local Similarity 94.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCACGACTGACGACGACC 20
|||||
Db 352 CCACGCTGACGACGACC 370

RESULT 15
AC021605 197764 bp DNA linear HTG 11-AUG-2001
LOCUS Homo sapiens clone RP11-225D16, WORKING DRAFT SEQUENCE, 32
DEFINITION unordered pieces.
AC021605
AC021605.4 GI:15148279
VERSION 1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 197764)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE JOURNAL
JOURNAL Homo sapiens, clone RP11-225D16
Unpublished
2 (bases 1 to 197764)
REFERENCE Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckwith,R., Bede,F.,
Bocuniaksky,L., Bouckhalter,B., Brown,A., Burtell,G., Castle,A.,
Choquel,Y., Colangelo,M., Collins,S., Collumore,A., Cooke,P.,
Darellano,K., Dewar,K., Domingo,M., Doyle,M., Fensholt,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heatord,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karasas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lien,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McKean,P., McKernan,K.,
McPheters,R., Meldrum,J., Menzies,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,J.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rotman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 11, 2001 this sequence version replaced g1:7767808.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5810
Center clone name: 225-D-16
----- Summary Statistics
Sequencing vector: M13; M77815; 49% of reads
Sequencing vector: Plasmid; n/a; 51% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 184156 bases at least Q40
Consensus quality: 190357 bases at least Q30
Consensus quality: 193088 bases at least Q20

```

*	1	13523:	contig of 13522 bp in length
*	13524	13623:	gap of 100 bp
*	13624	14403:	contig of 680 bp in length
*	14304	14403:	contig of 680 bp in length
*	14404	14505:	contig of 652 bp in length
*	15056	15155:	gap of 100 bp
*	15156	15778:	contig of 623 bp in length
*	15779	15878:	gap of 100 bp
*	15879	16229:	contig of 351 bp in length
*	16230	16329:	gap of 100 bp
*	16330	16687:	contig of 558 bp in length
*	16688	16987:	gap of 100 bp
*	16988	17233:	contig of 246 bp in length
*	17234	17333:	gap of 100 bp
*	17334	17999:	contig of 666 bp in length
*	18000	18099:	gap of 100 bp
*	18100	18777:	contig of 678 bp in length
*	18778	18877:	gap of 100 bp
*	18878	19460:	contig of 583 bp in length
*	19461	19560:	gap of 100 bp
*	19561	20742:	contig of 1182 bp in length
*	20743	20842:	gap of 100 bp
*	20843	21533:	contig of 691 bp in length
*	21534	21633:	gap of 100 bp
*	21634	22342:	contig of 709 bp in length
*	22343	22442:	gap of 100 bp
*	22443	23111:	contig of 669 bp in length
*	23112	23211:	gap of 100 bp
*	23212	23892:	contig of 681 bp in length
*	23893	23992:	gap of 100 bp
*	23993	24554:	contig of 862 bp in length
*	24555	24954:	gap of 100 bp
*	24955	26017:	contig of 1063 bp in length
*	26018	26117:	gap of 100 bp
*	26118	26328:	contig of 211 bp in length
*	26329	26428:	gap of 100 bp
*	26429	27272:	contig of 844 bp in length
*	27273	27372:	gap of 100 bp
*	27373	28285:	contig of 913 bp in length
*	28286	28385:	gap of 100 bp
*	28386	29253:	contig of 868 bp in length
*	29254	29353:	gap of 100 bp
*	29354	30221:	contig of 868 bp in length
*	30222	30321:	gap of 100 bp
*	30322	31453:	contig of 1132 bp in length
*	31454	31553:	gap of 100 bp
*	31554	33045:	contig of 1492 bp in length
*	33046	33145:	gap of 100 bp
*	33146	33347:	contig of 802 bp in length
*	33348	34047:	gap of 100 bp
*	34048	35747:	contig of 1700 bp in length
*	35748	35847:	gap of 100 bp
*	35848	36927:	contig of 1080 bp in length
*	36928	37027:	gap of 100 bp
*	37028	42128:	contig of 5101 bp in length
*	42129	42228:	gap of 100 bp
*	42229	50923:	contig of 8695 bp in length
*	50924	51023:	gap of 100 bp
*	51024	76663:	contig of 25840 bp in length
*	76664	76563:	gap of 100 bp
*	76564	123889:	contig of 46926 bp in length

FEATURES	*	123890	123989:	gap of	100	bp
source	*	123990	197764:	contig of	73775	bp in length
			Location/Qualifiers			
			1. .197764			
			/organism="Homo sapiens"			
			/db_xref="taxon:9606"			
			/clone="RP11-725D16"			
			/clone_lib="RCF11 Human Male BAC"			
			1. .13523			
misc_feature			/note="assembly-fragment"			
			clone_end:SP6			
			vector_side:left"			
misc_feature			13624. .14303			
			/note="assembly-fragment"			
misc_feature			14404. .15055			
			/note="assembly-fragment"			
misc_feature			15156. .15778			
			/note="assembly-fragment"			
misc_feature			15879. .16229			
			/note="assembly-fragment"			
misc_feature			16330. .16887			
			/note="assembly-fragment"			
misc_feature			16988. .17233			
			/note="assembly-fragment"			
misc_feature			17334. .17999			
			/note="assembly-fragment"			
misc_feature			19100. .18777			
			/note="assembly-fragment"			
misc_feature			18878. .19460			
			/note="assembly-fragment"			
misc_feature			19561. .20742			
			/note="assembly-fragment"			
misc_feature			20843. .21533			
			/note="assembly-fragment"			
misc_feature			21634. .22542			
			/note="assembly-fragment"			
misc_feature			22443. .23111			
			/note="assembly-fragment"			
misc_feature			23212. .23892			
			/note="assembly-fragment"			
misc_feature			23993. .24854			
			/note="assembly-fragment"			
misc_feature			24955. .26017			
			/note="assembly-fragment"			
misc_feature			26118. .26328			
			/note="assembly-fragment"			
misc_feature			26429. .27272			
			/note="assembly-fragment"			
misc_feature			27373. .28285			
			/note="assembly-fragment"			
misc_feature			28386. .29253			
			/note="assembly-fragment"			
misc_feature			29354. .30221			
			/note="assembly-fragment"			
misc_feature			30322. .31453			
			/note="assembly-fragment"			
misc_feature			31554. .33045			
			/note="assembly-fragment"			
misc_feature			33146. .33947			
			/note="assembly-fragment"			
misc_feature			34048. .35747			
			/note="assembly-fragment"			
misc_feature			35848. .36927			
			/note="assembly-fragment"			
misc_feature			37028. .42128			
			/note="assembly-fragment"			
misc_feature			43229. .50923			
			/note="assembly-fragment"			
misc_feature			51024. .76863			
			/note="assembly-fragment"			
misc_feature			76964. .123889			
			/note="assembly-fragment"			

Query Match 75.7%; Score 17.4; DB 2; Length 197764;
Best Local Similarity 94.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CCACGACTGAGCACCAC 20
|||||
DB 84705 CCACGCTGTGAGCAGCAC 84723

Search completed: October 17, 2002, 09:32:13
Job time : 1466.27 secs